



HAL
open science

Classification of hyperspectral data using spectral-spatial approaches

Yuliya Tarabalka

► **To cite this version:**

Yuliya Tarabalka. Classification of hyperspectral data using spectral-spatial approaches. Signal and Image processing. Institut National Polytechnique de Grenoble - INPG, 2010. English. NNT : . tel-00557734

HAL Id: tel-00557734

<https://theses.hal.science/tel-00557734>

Submitted on 19 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITY OF ICELAND
UNIVERSITE DE GRENOBLE
INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

THÈSE

pour obtenir le grade en co-tutelle de

DOCTEUR de l'Université de Grenoble
délivré par l'Institut polytechnique de Grenoble
Grenoble INP

Spécialité : « Signal Image Parole et Télécom »

préparée au Grenoble Images Parole Signal Automatique Laboratoire (Gipsa-Lab)
dans le cadre de l'École Doctorale « **Electronique , Electrotechnique, Automatique,
Traitement du Signal** »

and

PhD in Electrical Engineering

from the Faculty of Electrical and Computer Engineering, **University of Iceland**

présentée et soutenue publiquement

par

Yuliya TARABALKA

le 14/06/10

Titre :

CLASSIFICATION OF HYPERSPECTRAL DATA USING SPECTRAL-SPATIAL APPROACHES

Directeurs de thèse : Jón Atli BENEDIKTSSON et Jocelyn CHANUSSOT

JURY

Mme. Melba CRAWFORD	Président
M. Jean-Yves TOURNERET	Rapporteur
Mme. Xiuping JIA	Rapporteuse
M. Antonio PLAZA	Rapporteur
M. Paolo GAMBA	Rapporteur
M. Johannes R. SVEINSSON	Examineur
M. Jón Atli BENEDIKTSSON	Directeur de thèse
M. Jocelyn CHANUSSOT	Directeur de thèse

Abstract

Y. Tarabalka. *Classification of Hyperspectral Data Using Spectral-Spatial Approaches*. Ph.D. thesis, University of Iceland and Grenoble Institute of Technology, 2010.

Hyperspectral imaging records a detailed spectrum of the received light in each spatial position in the image. Since different substances exhibit different spectral signatures, hyperspectral imagery is a well-suited technology for accurate image classification, which is an important task in many application domains. However, the high dimensionality of the data presents challenges for image analysis. While most of the previously proposed classification techniques process each pixel independently without considering information about spatial structures, recent research in image processing has highlighted the importance of the incorporation of spatial context in a classifier.

In this thesis, we propose and develop novel spectral-spatial methods and algorithms for accurate classification of hyperspectral data. First, the integration of the Support Vector Machines (SVM) technique within a Markov Random Fields (MRFs) framework for context classification is investigated. SVM and MRF models are two powerful tools for high-dimensional data classification and for contextual image analysis, respectively.

In a second step, we propose classification methods using adaptive spatial neighborhoods derived from region segmentation results. Different segmentation techniques are investigated and extended to the case of hyperspectral images. Then, approaches for combining the extracted spatial regions with spectral information in a classifier are developed.

In a third step, we concentrate on approaches to reduce oversegmentation in an image, which is achieved by automatically “marking” the spatial structures of interest before performing a marker-controlled segmentation. Our proposal is to analyze probabilistic classification results for selecting the most reliably classified pixels as markers of spatial regions. Several marker selection methods are proposed, using either individual classifiers, or a multiple classifier system. Then, different approaches for marker-controlled region growing are developed, using either watershed or Minimum Spanning Forest methods and resulting in both segmentation and context classification maps.

Finally, we explore possibilities of high-performance parallel computing on commodity processors for reducing computational loads. The new techniques, developed in this thesis, improve classification results, when compared to previously proposed methods, and thus show great potential for various image analysis scenarios.

Key words: hyperspectral data, high-dimensional image processing, spectral-spatial classification, segmentation, support vector machines, Markov Random Field, multiple classifiers, minimum spanning forest, high-performance parallel computing.

Résumé

L'imagerie hyperspectrale enregistre un spectre détaillé de la lumière reçue dans chaque position spatiale de l'image. Comme des matières différentes manifestent des signatures spectrales différentes, l'imagerie hyperspectrale est une technologie bien adaptée pour la classification précise des images, ce qui est une tâche importante dans beaucoup de domaines appliqués. Cependant, la grande dimension des données complique l'analyse des images. La plupart des techniques de classification proposées précédemment traitent chaque pixel indépendamment, sans considérer l'information sur les structures spatiales. Cependant, la recherche récente en traitement d'images a souligné l'importance de l'incorporation du contexte spatial dans les classifieurs.

Dans cette thèse, nous proposons et développons des nouvelles méthodes et algorithmes spectro-spatiaux pour la classification précise des données hyperspectrales. D'abord, l'intégration de la technique des Machines à Vecteurs de Support (SVM) dans le cadre des Champs Aléatoires de Markov (MRFs) pour la classification contextuelle est étudiée. Les SVM et les modèles markoviens sont les deux outils efficaces pour la classification des données de grande dimension et pour l'analyse contextuelle d'images, respectivement.

Dans un second temps, nous avons proposé des méthodes de classification qui utilisent des voisinages spatiaux adaptatifs dérivés des résultats d'une segmentation. Nous avons étudié différentes techniques de segmentation et nous les avons adaptées pour le traitement des images hyperspectrales. Ensuite, nous avons développé des approches pour combiner les régions spatiales avec l'information spectrale dans un classifieur.

Nous avons aussi étudié des techniques pour réduire la sur-segmentation en utilisant des marqueurs des structures spatiales d'intérêt afin d'effectuer la segmentation par marqueurs. Notre proposition est d'analyser les résultats de la classification probabiliste afin de sélectionner les pixels les plus fiablement classés comme des marqueurs des régions spatiales. Nous avons proposé plusieurs méthodes pour la sélection des marqueurs, qui utilisent soit des classifieurs individuels, soit un ensemble de classifieurs. Ensuite, nous avons développé des techniques pour la segmentation par croissance de régions issues des marqueurs, en utilisant soit la ligne de partage d'eaux, soit une forêt couvrante de poids minimal, qui ont pour résultat les cartes de segmentation et de classification contextuelle.

Finalement, nous considérons les possibilités du calcul parallèle à haute performance sur les processeurs d'un usage commode afin de réduire la charge du calcul. Les nouvelles méthodes développées dans cette thèse améliorent les résultats de classification par rapport aux méthodes proposées précédemment, et ainsi montrent un grand potentiel pour les différents scénarios de l'analyse d'image.

Mot clés: données hyperspectrales, traitement d'images de hautes dimensions, classification spectro-spatiale, segmentation, machines à vecteurs de support, champ aléatoire de Markov, ensemble de classifieurs, forêt couvrante de poids minimal, calcul parallèle à haute performance.

Ágrip

„Flokkun fjölrásagagna með aðferðum sem byggjast á róf- og rúmupplýsingum“. Doktorsritgerð. Sameiginleg prófgráða á milli Háskóla Íslands og INP Grenoble (INPG) háskólans í Grenoble, Frakklandi.

Markmið verkefnisins var að þróa aðferðir sem nota ekki aðeins klassískar rófupplýsingar við flokkun fjarkönnunarmynda heldur einnig rúmupplýsingar úr myndunum sjálfum.

Fjarkönnunarmyndir með mikilli rófupplausn geyma nákvæmar upplýsingar um yfirborð jarðar. Myndirnar geyma mismunandi rófsvörun frá mismunandi yfirborðsflokkum ásamt því að sýna stafrænar myndir af yfirborðinu. Svona fjarkönnunarmyndefni er vel til þess fallið að flokka yfirborð jarðarinnar á nákvæman hátt. Hins vegar er flokkun á svona myndum ekki einfalt verkefni þar sem myndirnar geyma rófupplýsingar af hárrí vidd, vidd sem hleypur að jafnaði á hundruðum. Umfang gagnanna skapar því ýmis vandamál við flokkunina. Flestar aðferðir sem settar hafa verið fram til flokkunar myndefnis með hárrí rófupplausn nýta aðeins rófupplýsingarnar, og flokka hverja einstaka mynddeiningu án þess að skoða næstu mynddeiningar og hvaða rúmupplýsingar myndin sjálf hefur að geyma.

Í doktorsritgerðinni eru þróaðar nýjar aðferðir sem nýta bæði róf- og rúmupplýsingarnar í fjarkönnunarmyndum til nákvæmnar flokkunar gagna með mikilli rófupplausn. Í ritgerðinni eru settar fram nokkrar slíkar aðferðir. Ein þeirra byggir á samþættingu stoðviðravéla (e. Support Vector Machines, SVM) og Markov slembisviða (e. Markov Random Field, MRF). SVM er öflug aðferð til flokkunar rófupplýsinga, en MRF er notað til að meta samhengið í myndefninu. Önnur aðferð sem sett er fram í ritgerðinni byggir á notkun aðhæfðs nágrennis í rúmi (þ.e. innan myndefnis), en aðhæfða nágrennið byggir á bútun (e. segmentation) svæða í myndinni. Nokkrar bútunaraðferðir eru rannsakaðar og útvíkkaðar fyrir fjarkönnunarmyndefni af hárrí rófupplausn. Rúmupplýsingar sem fást með þessari aðferð eru síðan notaðar ásamt rófupplýsingum við flokkun. Þriðja aðferðin sem sett er fram kemur í veg fyrir ofmikla bútun (oversegmentation) í myndefni. Svokallaðri mörkun (e. marking) er beitt til að ná þessu markmiði. Nokkrar mismunandi aðferðir við mörkun eru settar fram fyrir hefðbundna flokkara annars vegar og margflokka (e. multiple classifiers) hins vegar. Síðan eru aðferðir sem byggja á vatnaskilsbútun (e. watershed segmentation) eða lágmarks spönnunarskógi (e. Minimum Spanning Forest, MSF) notaðar við ræktun svæða. Niðurstaðan verða kort með bútun svæðanna og flokkunarkort sem byggja á samhengi. Í síðasta hluta ritgerðarinnar er hugað að því hvernig lágmarka megi reikniþunga með notkun samhliðareikninga er nota örgjörva sem hafa mikila afkastagetu.

Aðferðirnar sem settar eru fram í ritgerðinni eru mikilvægt framlag til flokkunar á fjarkönnunarmyndum með hárrí róf- og rúmupplausn. Nákvæmni niðurstaðnanna meiri en fengist hefur áður með sömu gögn og gefur góðfyrirheit um notkun aðferðanna í mismunandi notkun.

Lykilorð: Fjarkönnunargöng með mikilli rófupplausn, myndvinnsla í hárrí vidd, flokkun gagna bæði í rófi og rúmi, bútun, stoðvigravélar, Markov slembisvið, margflokka, lágmarks spönnuar skógur, afkastamiklir samhliðareikningar.

Acknowledgments

First and foremost I offer my sincerest gratitude to my supervisors, Jocelyn Chanussot and Jón Atli Benediktsson, who have supported me throughout my thesis with their patience and knowledge whilst allowing me the room to work in my own way. One simply could not wish for better or friendlier supervisors. A special thanks to Jocelyn Chanussot, who has passionated me by image processing and analysis subject during my Master studies and who has involved me in this exciting adventure of working on my thesis in Norway, Iceland and France in the highly international environment. I thank Jocelyn for providing me the opportunity to expand my horizon and vision of our beautiful world, to meet many new cultures and learn many new things on our planet and beyond.

I thank Torbjørn Skauli, Trym and Ingebjørg for helping me to acquire first knowledge on hyperspectral imaging and remote sensing and for taking such good care during my stay in Kjeller, Norway. I would like to thank Taras Rafa, Mykhailo Dychkovskyy and Mykhailo Fryz, my first research advisors, for all the discussions we had and for inspiring me to pursue my research investigations. I am very grateful to Pierre Badin, Gérard and Frédéric for their effort, encouragement and collaboration during my Master studies in Grenoble, France. In many ways I have learnt much from and because of them.

I thank Melba Crawford for accepting to preside over my thesis committee and for her interest and suggestions regarding my work. I am grateful to Jean-Yves Tournet, Xiuping Jia, Antonio Plaza and Paolo Gamba for their comments and suggestions regarding my manuscript and for attending my PhD defense in Iceland. I would like to thank Johannes Sveinsson for interesting discussions during my PhD work. I thank Jim Tilton and other collaborators for their help and valuable suggestions.

I thank my parents for giving me life, which is so much wonderful and passionate, and for their love, help and support. I thank John, my best friend, for his patience and encouragement when it was most required, and for all the unforgettable time together. I am grateful to my colleagues for providing me a pleasant and creative environment during my research. Finally, I would like to thank all interesting people from all over the world who have contributed directly or indirectly to the realisation of this project, and who have made my time in Norway, Iceland and France enjoyable and fruitful.

Many thanks! Merci beaucoup! Takk fyrir! Tusen takk!

Yuliya

List of Abbreviations

AA	Average Accuracy
API	Application Programming Interface
ARCHER	Airborne Real-time Cueing Hyperspectral Enhanced Reconnaissance
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
CA	Class-specific Accuracy
CC	Connected Components
CMG	Color Morphological Gradient
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture
DAFE	Discriminant Analysis Feature Extraction
DBFE	Decision Boundary Feature Extraction
DMP	Derivative Morphological Profile
ECHO	Extraction and Classification of Homogeneous Objects
EM	Expectation Maximization
EMP	Extended Morphological Profile
FNEA	Fractal Net Evolution Approach
GLCM	Gray Level Co-occurrence Matrix
GPU	Graphics Processing Unit
HSEG	Hierarchical SEGmentation
HSWO	Hierarchical Step-Wise Optimization
HYDICE	HYperspectral Digital Imagery Collection Experiment
ICA	Independent Component Analysis
ISODATA	Iterative Self-Organizing DATA Analysis
LDA	Linear Discriminant Analysis
MAP	Maximum A-Posteriori
MC	Multiple Classifier
ML	Maximum Likelihood
MNF	Maximum Noise Fraction
MP	Morphological Profile
MRF	Markov Random Field
MS	Mean Shift

MSF	Minimum Spanning Forest
MSSC	Multiple Spectral-Spatial Classifier
NMF	Non-negative Matrix Factorization
NN	Nearest Neighbor
NWFE	Non-parametric Weighted Feature Extraction
OA	Overall Accuracy
PC	Principal Component
PCA	Principal Component Analysis
PCFA	Piecewise Constant Function Approximations
PDF	Probability Density Function
RBF	Radial Basis Function
RCMG	Robust Color Morphological Gradient
RHSEG	Recursive Hierarchical SEGmentation
RODIS	Reflective Optics System Imaging Spectrometer
SAR	Synthetic Aperture Radar
SAM	Spectral Angle Mapper
SDK	Software Development Kit
SE	Structuring Element
SID	Spectral Information Divergence
SVM	Support Vector Machines

Contents

I	Research Summary	1
1	Introduction	3
1.1	Context of the Research Work	3
1.2	Background	5
1.2.1	Pixelwise classification	5
1.2.2	Spectral-spatial classification	10
1.3	Objectives	17
1.4	Main Contributions	17
1.5	Thesis Outline	19
2	Proposed Methods	21
2.1	Context Classification Using Closest Fixed Neighborhoods, SVM and MRFs	21
2.1.1	Probabilistic SVM classification	21
2.1.2	Computation of the gradient	22
2.1.3	MRF-based regularization	23
2.1.4	Concluding discussion	24
2.2	Spectral-spatial Classification Using Adaptive Neighborhoods	24
2.2.1	Spatial-based segmentation using watershed transformation	24
2.2.2	Spectral-based segmentation using partitional clustering	28
2.2.3	Hierarchical segmentation both in the spatial and the spectral domain	31
2.2.4	Spectral-spatial classification approaches	31
2.2.5	Concluding discussion	33
2.3	Segmentation and Classification Using Automatically Selected Markers	34
2.3.1	Marker selection	34
2.3.2	Classification using marker-controlled region growing	40
2.3.3	Concluding discussion	44
2.4	Experimental Evaluation	45
2.4.1	Data sets	46
2.4.2	Experimental results	50
2.4.3	Concluding discussion	64
2.5	Implementation Issues	65
2.5.1	Consideration of the computational cost for the proposed classification methods . .	65
2.5.2	Parallel GPU processing of hyperspectral imagery	67
3	Conclusions	69
3.1	Contributions	69
3.2	Perspectives	70
3.3	List of Publications	71

Résumé étendu	73
R.1 Introduction	73
R.2 État de l'Art	74
R.2.1 Classification au niveau pixel	74
R.2.2 Classification spectro-spatiale	76
R.3 Objectifs	81
R.4 Contributions Principales	81
R.5 Plan de la Thèse	83
R.6 Évaluation expérimentale	84
R.7 Conclusions et Perspectives	88
R.7.1 Contributions	89
R.7.2 Perspectives	90
Bibliography	92
II Publications	103
Paper 1. Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," <i>IEEE Geoscience and Remote Sensing Letters</i> , 2010, DOI 10.1109/LGRS.2010.2047711.	105
Paper 2. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," <i>Pattern Recognition</i> , vol. 43, no. 7, pp. 2367-2379, July 2010.	113
Paper 3. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," <i>IEEE Trans. on Geoscience and Remote Sensing</i> , vol. 47, no. 8, pp. 2973-2987, Aug. 2009.	129
Paper 4. Y. Tarabalka, J. Chanussot, J. A. Benediktsson, "Classification based marker selection for watershed transform of hyperspectral images," in Proc. of <i>IGARSS'09</i> , Cape Town, South Africa, 2009, pp. III-105 - III-108.	147
Paper 5. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using Minimum Spanning Forest grown from automatically selected markers," <i>IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics</i> , 2010, DOI 10.1109/TSMCB.2009.2037132.	153
Paper 6. Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Multiple spectral-spatial classification approach for hyperspectral data," <i>IEEE Trans. on Geoscience and Remote Sensing</i> , under review.	169
Paper 7. Y. Tarabalka, T. V. Haavardsholm, I. Kasen, and T. Skauli, "Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing," <i>Journal of Real-Time Image Processing</i> , vol. 4, no. 3, pp. 287-300, Aug. 2009.	183

Part I

Research Summary

Chapter 1

Introduction

THIS chapter presents a general framework of this thesis. First, the context of the research work is introduced and the necessary background material is reviewed. Then, the thesis objectives and main contributions are described. These contributions are further discussed in Chapter 2, which gives an extended summary of the new methods and algorithms for hyperspectral data classification, and in Part II, which is composed of seven publications presenting the main findings of this thesis.

1.1 Context of the Research Work

The concept of hyperspectral imaging, also known as imaging spectroscopy, originated in the 1980's, when A. F. H. Goetz and his colleagues at NASA's Jet Propulsion Laboratory began a revolution in remote sensing by developing new optical instruments such as the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [59, 111]. *Imaging spectrometry* is defined as "the simultaneous acquisition of images in many narrow, contiguous spectral bands" [54]. Thus, *hyperspectral imaging* is concerned with the measurement, processing and analysis of spectra acquired from a given scene at a short, medium or long distance by typically an airborne or satellite sensor. Examples of hyperspectral airborne imaging systems are AVIRIS [59], HYperspectral Digital Imagery Collection Experiment (HYDICE) [118], Reflective Optics System Imaging Spectrometer (ROSIS) [50], Airborne Real-time Cueing Hyperspectral Enhanced Reconnaissance (ARCHER) [131], HyMap [27]. The hyperspectral sensors currently operating in space are Hyperion (USA, 2000) [108], Compact High Resolution Imaging Spectrometer (CHRIS, UK, 2001), HySI (India, 2008), HJ-1A (China, 2008), Advanced Responsive Tactically Effective Military Imaging Spectrometer (ARTEMIS, USA, 2009) [129]. New sensors/missions are under development, such as Multi-Sensor Microsatellite Imager (MSMI, South Africa, 2010), Hyperspectral Precursor of the Application Mission (PRISMA, Italy, 2012-2013), Environmental Mapping Program (EnMAP, Germany, 2014) [129]. The hyperspectral sensors typically cover a range of 0.4 to 2.5 μm using 115 to 512 spectral channels, with a spatial resolution varying from 0.75 to 20 m/pixel for airborne sensors and from 5 to 506 m/pixel for satellite sensors.

Figure 1.1 illustrates the hyperspectral imaging concept. As can be seen from the figure, every pixel can be represented as a high-dimensional vector across the wavelength dimension containing the sampled reflectance spectrum. Since different substances exhibit different spectral signatures, hyperspectral imaging is a well-suited technology for numerous remote sensing applications including accurate image classification. *Hyperspectral image classification*, which can be defined as identification of objects in a scene captured by a hyperspectral imaging sensor, is an important task in many application domains such as [82]:

- Precision agriculture: Hyperspectral remote sensing is extensively used for monitoring the development and health of crops and thus preventing disease outbreaks.
- Mineralogy: Many minerals can be identified from hyperspectral images, and their relation to the

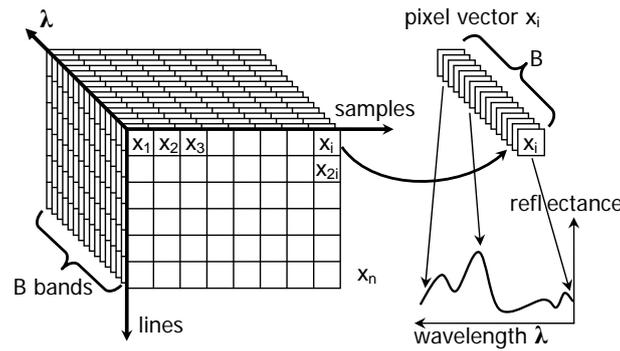


Figure 1.1: Schematic illustration of the hyperspectral imaging concept. Images are acquired simultaneously of up to several hundred narrow spectral bands providing a detailed reflectance spectrum for every pixel in the scene.

presence of valuable minerals is often understood. Currently the researchers investigate the effect of oil and gas leakages from pipelines and natural wells on the spectral signatures of vegetation.

- Monitoring and management of environment: Classification algorithms can be used for creating thematic maps for analyzing the urban area growth and sensitivity of the different city areas to natural risks, as well as investigating changes in the land and coastal-ocean ecosystems.
- Surveillance: Hyperspectral imaging is particularly useful in military surveillance, for detecting objects that are hidden from the naked eye and thus providing security services.

However, a large number of spectral channels implies high dimensionality of the data and presents challenges to image analysis and classification. First attempts to analyze hyperspectral images were based on existing techniques developed for multispectral images (composed of a few, usually three to seven, spectral channels). However, most of the commonly used methods designed for the analysis of gray level, color or multispectral images are not appropriate for hyperspectral images. As a matter of fact, very limited reference data¹ are usually available, which limits the performances of supervised classification techniques. Furthermore, for analysis of hyperspectral images the well known curse of dimensionality prevents robust statistical estimations (this problem is known as the Hughes phenomenon [66]). Therefore, to take full advantage of the rich information provided by the spectral dimension, the development of new algorithms is required.

In the last decade, a lot of new developments for hyperspectral image classification have been proposed and investigated. This proves the increasing interest in analysis of hyperspectral data. Most classification methods process each pixel independently without considering the correlations between spatially adjacent pixels (so called *pixelwise* classifiers) [20, 53, 81]. In particular, Support Vector Machines (SVM) have shown good performances in terms of classification accuracies for classifying high-dimensional data when a limited number of training samples is available [20, 43, 148].

However, if pixelwise classification enables the precise characterization of the sensed materials, it does not take contextual information into account. If a random permutation of all the pixels within a picture is performed, the classification of each value is not affected, while our visual understanding and interpretation of the scene is seriously harmed (one only sees “noise” after the global random permutation). This toy experiment underlines the importance of spatial contextual information. Consequently, in order to improve classification results, it is very important to develop *spectral-spatial* classification techniques capable to consider spatial dependencies between pixels [43, 45, 74, 116, 138]. One of the spectral-spatial classification approaches consists in including the information from the closest neighborhood when clas-

¹By *reference data*, we mean manually labeled pixels which are used for training classifiers followed by assessment of classification accuracies.

sifying each pixel. Markov Random Fields (MRFs) [42, 67], morphological filtering [43, 45, 109] and composite kernels [21, 22, 78] have been investigated for this purpose. Although classification accuracies have been improved when compared to pixelwise classification, the use of these methods raises the problem of scale² selection.

The general aim of this thesis is to further develop methods for accurate classification of hyperspectral data using both the spectral and the spatial information. The focus of the thesis is the incorporation of spatial contextual information into the classification procedure. This raises two principal questions: 1) How to define spatial structures, or neighborhoods for each pixel in the image automatically? 2) How to combine spectral and spatial information in classification? In the following, we will review the necessary background material for classification of hyperspectral images. Then, we will precise the objectives of this thesis work.

1.2 Background

1.2.1 Pixelwise classification

The general hyperspectral image *classification problem* can be described as follows: At the input a B -band hyperspectral cube is given, which can be considered as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. Let $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ be a set of information classes in the scene. Classification (*supervised classification*) consists in assigning each pixel to one of the K classes of interest.

An information class can represent either a physical substance (a ground cover material, for instance, *snow, water, wheat*), or a specific group of objects which may be made of several different physical materials (for instance, *roof, shadows, trees*). In the first case, one may believe that each ground cover material, and thus each class, can be appropriately represented by a single spectral curve, called *spectral signature* of this material. However, the angle of the sun, the direction of view, the atmospheric condition and a number of other such uncontrollable variables substantially affect the acquired spectral response of any given material [80]. Since it is a daunting problem (not resolved yet) to compensate these variables, the spectral response of any given class varies significantly over the image scene. It is evident that in the case when a class represents several ground cover materials, variations of the spectral response within this class are even more significant. The important rule for establishing the list of classes is that the classes must be separable in terms of available spectral features.

David Landgrebe and his coworkers from the Laboratory for Applications of Remote Sensing at Purdue University were seminal in exploring procedures for hyperspectral data analysis and classification. They adapted pattern recognition procedures for this purpose, and their proposed classification scheme (depicted in Figure 1.2) is now widely used for thematic mapping from hyperspectral images [80, 81, 116].

The basic idea of the *pixelwise* pattern recognition approach is as follows: Each image pixel is considered as a pattern and its spectrum is considered as the initial set of features. Since this set of features is often redundant, feature extraction/selection steps are performed aiming at reducing the dimensionality of the feature set (from B dimensions in the original data to D dimensions in a new feature space) and maximizing separability between classes. Different feature extraction techniques have been proposed and explored for this purpose, such as Discriminant Analysis Feature Extraction (DAFE), Decision Boundary Feature Extraction (DBFE), Non-parametric Weighted Feature Extraction (NWFE) [43, 81]. The next step (called “*training of the classifier*”) consists in partitioning the entire feature space into K exhaustive, nonoverlapping regions, so that every point in the feature space is uniquely associated with one of the K classes. Once this step is accomplished, each pixel is classified according to its feature set.

In order to partition the feature space into decision regions, a set of representative samples for each class, referred to as *training samples*, is used. It is usually obtained by manually labeling a small number of pixels in an image. Figure 1.3 illustrates an example of the training samples set for a ROSIS airborne image. The training samples are further used for defining a model of the classes in the feature space.

²By *scale*, we mostly refer to the size of spatial objects of interest.

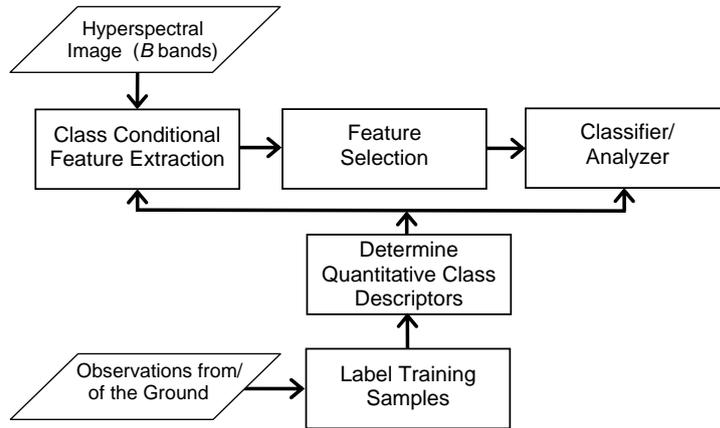


Figure 1.2: Schematic diagram of the hyperspectral data classification and analysis process (proposed by Landgrebe [80]).

Landgrebe has proposed to use signal based processing algorithms for this purpose. Each class ω_i can be modeled in terms of a class-conditional Probability Density Function (PDF) $p(\mathbf{x}|\omega_i)$. Assuming that each class can be described by a Gaussian (normal) distribution [40], Gaussian Maximum Likelihood (ML) classification has been for many years the standard thematic mapping procedure in hyperspectral remote sensing [116]. Essentially, it assigns a given pixel to the class ω_i maximizing the posterior probability $p(\omega_i|\mathbf{x})$. By using the Bayes rule, this is equivalent to maximizing

$$p(\mathbf{x}|\omega_i)p(\omega_i), \quad (1.1)$$

where $p(\omega_i)$ is the prior probability of the class ω_i . The Gaussian class-conditional PDF is defined as

$$p(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x} - \boldsymbol{\mu}_i)\right\}, \quad (1.2)$$

where D is the dimensionality of the feature space, $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean vector and covariance matrix for the class ω_i that can be estimated by

$$\hat{\boldsymbol{\mu}}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \mathbf{x}_{j,i}, \quad (1.3)$$

$$\hat{\boldsymbol{\Sigma}}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} (\mathbf{x}_{j,i} - \boldsymbol{\mu}_i)(\mathbf{x}_{j,i} - \boldsymbol{\mu}_i)^T, \quad (1.4)$$

where $\mathbf{x}_{j,i}, j = 1, \dots, m_i$ are the training samples for the class ω_i .

A serious drawback of this method consists in the primary assumption about the shape of the class-conditional PDFs. If this assumption is wrong, classification results are no more accurate. Furthermore, the high number of features available, usually coupled with the limited number of training samples, prevents reliable estimation of statistical class parameters. As a result, with a limited training set, the classification accuracy tends to decrease as the dimensionality increases [80, 111]. High-dimensional spaces are mostly empty [72], making density estimation even more difficult.

A simpler ML classification consists in modeling each class only by its mean response (the classes are assumed to have unit variance in all features and the features are assumed to be uncorrelated to each other). This method, called *minimum distance to means* classifier, assigns each pixel to the class with the

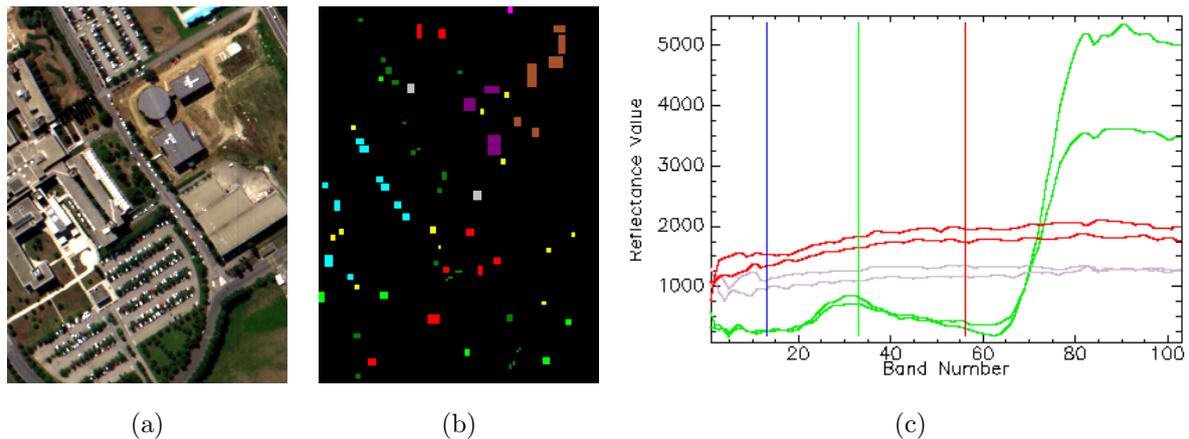


Figure 1.3: ROSIS data. (a) Three-band color composite (650, 558, and 478 nm). (b) Set of training samples. Nine information classes are considered: Asphalt, Meadows, Gravel, Trees, Metal sheets, Bare soil, Bitumen, Bricks and Shadows. (c) A display of the data of six pixels belonging to three classes *asphalt*, *meadows* and *bricks* as a function of spectral band number (wavelength). Graph colors correspond to the colors of classes. Both inter-class and intra-class variability can be observed.

closest mean vector [80]. An appropriate distance measure must be chosen. The Spectral Angle Mapper (SAM) algorithm, using the SAM distance for computing the similarity of spectra in multidimensional space, is one of the simplest supervised classification approaches for hyperspectral data [76].

Definition 1.1 (SAM measure) *The SAM measure determines the spectral similarity between two vectors $\mathbf{x}_i = (x_{i1}, \dots, x_{iB})^T$ and $\mathbf{x}_j = (x_{j1}, \dots, x_{jB})^T$ by computing the angle between them. It is defined as*

$$SAM(\mathbf{x}_i, \mathbf{x}_j) = \arccos \left(\frac{\sum_{b=1}^B x_{ib}x_{jb}}{[\sum_{b=1}^B x_{ib}^2]^{1/2} [\sum_{b=1}^B x_{jb}^2]^{1/2}} \right). \quad (1.5)$$

An important property of the SAM measure is that poorly illuminated and more brightly illuminated pixels of the same physical substance will be mapped to the same spectral angle despite their difference in illumination. The SAM method works well only in the case when intra-class spectra variability is low. In other cases, the classes cannot be accurately described by only their mean vectors, and SAM classification fails.

In the 1990s, *neural network* approaches for classifying hyperspectral images received a lot of attention in classification [14, 95, 133, 154]. Neural network models have an advantage over the statistical methods that they are distribution free and thus no prior knowledge about the statistical distribution of classes is needed. A set of weights and non-linearities describe the neural network, and these weights are computed in an iterative training procedure. The interest for such approaches greatly increased in the 1990s because of recently proposed feasible training techniques for nonlinearly separable data [13]. Yet the use of neural networks for hyperspectral image classification has been limited primarily due to their algorithmic and training complexity [116]. Vaiphasa [147] has proposed genetic algorithms for classification of hyperspectral data, capable to deal with nonlinearly separable patterns, but computationally demanding.

In recent years, SVM and the use of kernels to transform data into a new feature space where linear separability can be exploited have been proposed and have shown to be extremely well suited to classify high-dimensional data when a limited number of training samples is available [20, 43, 60, 148]. The SVM method attempts to separate training samples belonging to different classes by tracing maximum margin hyperplanes in the space where the samples are mapped. In the following, we briefly describe the general mathematical formulation of SVM for binary classification problems.

SVM classification

Let us consider a binary classification problem in a B -dimensional space \mathbb{R}^B , with N training samples, $\mathbf{x}_i \in \mathbb{R}^B$, and their corresponding labels $y_i = \pm 1$: $\{(\mathbf{x}_i, y_i) \mid i \in [1, N]\}$ available. The SVM method consists in finding the hyperplane that maximizes the margin (see Figure 1.4), i.e, the distance to the closest training data points in both classes. Noting $\mathbf{w} \in \mathbb{R}^B$ as the vector normal to the hyperplane and $b \in \mathbb{R}$ as the bias, the hyperplane H_p is defined as

$$\mathbf{w} \cdot \mathbf{x} + b = 0, \quad \forall \mathbf{x} \in H_p. \quad (1.6)$$

If $\mathbf{x} \notin H_p$ then

$$f(\mathbf{x}) = \frac{|\mathbf{w} \cdot \mathbf{x} + b|}{\|\mathbf{w}\|} \quad (1.7)$$

defines the distance of \mathbf{x} to H_p . In the linearly separable case, such a hyperplane has to satisfy:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1, \quad \forall i \in [1, N]. \quad (1.8)$$

One can always choose the vector \mathbf{w} after appropriate normalization such that condition (1.8) is satisfied. The optimal hyperplane is the one that maximizes the margin: $2/\|\mathbf{w}\|$. This is equivalent to minimizing $\|\mathbf{w}\|/2$ and leads to the following quadratic optimization problem:

$$\min \left[\frac{\|\mathbf{w}\|^2}{2} \right], \quad \text{subject to (1.8)}. \quad (1.9)$$

For taking into account non-linearly separable data, *slack* variables ξ are introduced to deal with misclassified samples (see Fig. 1.4). Eq. (1.8) becomes

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1 - \xi_i, \quad \xi_i \geq 0 \quad \forall i \in [1, N]. \quad (1.10)$$

The final optimization problem can be described as

$$\min \left[\frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^N \xi_i \right], \quad \text{subject to (1.10)} \quad (1.11)$$

where the constant C is a regularization parameter that controls the amount of penalty. These optimization problems are usually solved by quadratic programming [148]. The classification is further performed by computing $y_u = \text{sgn}(\mathbf{w} \cdot \mathbf{x}_u + b)$, where (\mathbf{w}, b) are the hyperplane parameters found during the training process and \mathbf{x}_u is an unseen sample.

An important notice is that the pixel vectors in the optimization and decision rule equations always appear in pairs related through a scalar product. Therefore, these products can be replaced by nonlinear functions of the pairs of vectors, essentially projecting the pixel vectors into a higher dimensional space \mathbb{H} and thus improving linear separability of data:

$$\begin{aligned} \mathbb{R}^B &\rightarrow \mathbb{H} \\ \mathbf{x} &\rightarrow \Phi(\mathbf{x}) \\ \mathbf{x}_i \cdot \mathbf{x}_j &\rightarrow \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j). \end{aligned} \quad (1.12)$$

Here, $\Phi(\cdot)$ is a nonlinear function for projecting feature vectors into a new space, $K(\cdot)$ is a *kernel* function that allows one to avoid the computation of the scalar products in the transformed space $[\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)]$ and thus reduces the computational complexity of the algorithm. The kernel K should fulfill Mercer's condition [19]. SVM require to choose the kernel carefully. For hyperspectral image classification, two kernels are widely used: the polynomial function and the Gaussian Radial Basis Function (RBF).

$$K_{POLY}(\mathbf{x}_i, \mathbf{x}_j) = [(\mathbf{x}_i \cdot \mathbf{x}_j) + 1]^p, \quad (1.13)$$

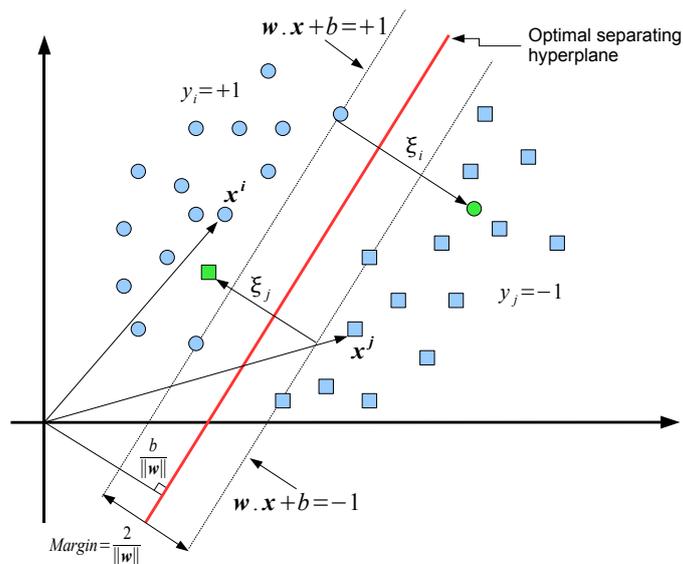


Figure 1.4: Classification of a non-linearly separable case by SVM. There is one non separable feature vector in each class.

$$K_{GAUSS}(\mathbf{x}_i, \mathbf{x}_j) = \exp \left[-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right], \quad (1.14)$$

where p is the order of the polynomial kernel function, and γ is the spread of the RBF kernel.

SVM are primarily designed to solve binary problems where the class labels can only take two values: ± 1 . To solve the K -class problem, various approaches have been proposed. Two main approaches [120], combining a set of binary classifiers, are defined as

- **One Versus All:** K binary classifiers are applied on each class against the others. Each pixel \mathbf{x}_i is assigned to the class with the *maximum* output $f(\mathbf{x}_i)$.
- **One Versus One:** $\frac{K(K-1)}{2}$ binary classifiers are applied on each pair of classes. Each pixel is assigned to the class getting the highest number of votes (*i.e.*, winning the *maximum* number of binary classification procedures).

The *one versus one* classification has shown to be the most suitable for large problems [64], and its computational complexity is comparable to that of the *one versus all* approach. Other multiclass strategies include decision trees and global training [94]. However, they have not demonstrated higher performances than the two described techniques.

A detailed analysis of the SVM theory and its value for hyperspectral image classification can be found in [19, 60, 94, 148]. As a conclusion, SVM directly exploit the geometrical properties of data, without involving a density estimation procedure. This method has proven to be much more effective than other nonparametric classifiers (such as neural networks and the k-Nearest Neighbor (k-NN) classifier [40]) in terms of classification accuracies, computational complexity and robustness to parameter setting. SVM can efficiently handle high-dimensional data, exhibiting low sensitivity to the Hughes phenomenon [66, 94]. Therefore, it is an excellent approach to avoid the usually time-consuming feature extraction/selection procedure and thus simplify the traditional pattern recognition scheme proposed by Landgrebe [80] and illustrated in Figure 1.2. Finally, it exhibits good generalization capability, fully exploiting the discrimination capability of the relatively few training samples available. All these advantages of the SVM method have made it the most widely used classifier for hyperspectral data in the last decade [86].

1.2.2 Spectral-spatial classification

The methods described so far are often referred as *pixelwise classifiers*, since they assign each pixel to one of the classes based on its spectral properties alone, with no account taken of how spatially adjacent pixels are classified. However, in a real image, neighboring pixels are related or correlated, both because imaging sensors acquire significant amount of energy from adjacent pixels [117] and because homogeneous structures occurred in the image scene are generally large compared to the size of a pixel. For instance, if a given pixel in the agricultural image scene represents a corn, its adjacent pixels belong with the high probability to the corn field. This spatial contextual information should help for an accurate scene interpretation. Therefore, in order to improve classification results, *spectral-spatial classification* methods, or *context classifiers*, must be developed, which assign each image pixel to one class based on: 1) its own spectral values (the spectral information) and 2) information extracted from its neighborhood (the spatial information). The use of context classifiers is especially rewarding when processing high resolution images and/or data with large spatial regions in the image scene. In the following, we describe the existing methods for spectral-spatial classification of hyperspectral data.

Landgrebe and his research team were pioneers in introducing spatial context into a multiband image classification. They have proposed the well-known ECHO (Extraction and Classification of Homogeneous Objects) classifier [74], which has been widely used by the remote sensing community and has become the standard technique for joint spectral/spatial classification. The method is based on region growing to find homogeneous groups of adjacent pixels that can be characterized by group means and covariances. These groups are further classified as single objects. Thus, the ECHO algorithm consists of two steps [74, 81]:

1. *Partitioning* into statistically homogeneous regions, using a two-stage conjunctive partitioning algorithm:
 - (a) First, all the pixels are divided by a rectangular grid into small groups, each group containing $b \times b$ pixels (parameter b must be set by the user). Each group becomes a unit called a “cell” if it satisfies a mild criterion of homogeneity. The threshold used for this test is also set by the user. Those groups that fail the test are referred to as “singular cells” and their individual pixels are classified [in step (2)] by an ML pixelwise classifier.
 - (b) At this stage, each non-singular cell is compared to an adjacent “field”, which is defined as a group of one or several connected cells that have previously been merged. The two considered samples are merged if they appear statistically similar by the chosen criterion. By successively merging adjacent cells if possible, an image partitioning into fields and individual pixels is obtained.
2. *Classification*: Each individual pixel and each field are classified by an ML classifier.

Since the ECHO classification involves estimation of covariance matrices, a feature extraction/ dimensionality reduction step is recommended before applying this technique to hyperspectral data (for reducing the number of parameters to be estimated). As a conclusion, the ECHO method has introduced a powerful and commonly used approach for spectral-spatial classification, which consists in partitioning the image into homogeneous regions and then classifying each region as a single object.

The last decade has witnessed major advances in context classification of hyperspectral images. A good survey of these developments can be found in [111]. The design of a spectral-spatial classification system implies two important steps:

1. An approach for *extracting spatial contextual information* from the image scene (*i.e.*, *neighborhood system*) must be defined. For instance, the ECHO classifier partitions an image into statistically homogeneous regions for this purpose.
2. A rule for *combining spectral and spatial information* in a classifier must be defined. Here, two categories of context classification methods can be distinguished: (a) Spatial contextual information is exploited in the classification stage. For instance, spectral and spatial information can be combined within a feature vector of each pixel, and then pixelwise classification technique can be applied to the obtained set of vectors [22, 45]. Another group of methods from this category firstly defines the objects within the image scene, and then classifies each object [74]. (b) Spatial

dependencies are considered in the decision rule. An example is a pixelwise classification followed by spatial regularization of the classification map [16, 114, 117].

Thus, a pattern classification approach (or several approaches) must be commonly chosen (for instance, Bayesian method, neural network, Linear Discriminant Analysis (LDA), SVM) to be used in the spectral-spatial classification system.

Table 1.1: State of the art of spectral-spatial classification methods for hyperspectral data (“classif” means classification).

Neighborhood system	ML and Bayesian methods	SVM	Other methods
Closest fixed neighbors	<p>★ 2000: Pony <i>et al.</i> [114] (ML classif + MRF regularization)</p> <p>★ 2002: Jackson and Landgrebe [67] (DAFE + Bayesian classif + MRF regularization)</p> <p>★ 2006: Tsai <i>et al.</i> [145] (GLCM + PCA + ML classif)</p>	<p>★ 2005: Farag <i>et al.</i> [42] (SVM regression for density estimation + Bayesian classif + MRF regularization)</p> <p>★ 2006: Camps-Valls <i>et al.</i> [22] (Mean and standard deviation + SVM composite kernel classif)</p> <p>★ 2006: Huang and Zhang [65] (NMF + GLCM + SVM classif using stacked features)</p>	<p>★ 2004: Dell’Acqua <i>et al.</i> [35] (DAFE + fuzzy ARTMAP with spatial reclassification)</p>
Morphological and area filtering	<p>★ 2009: Collet <i>et al.</i> [28] (modeling of galaxy structures + Bayesian classif)</p>	<p>★ 2007: Fauvel [43] (PCA + EMP + DFFE or NWFE + SVM classif using stacked features)</p> <p>★ 2007: Fauvel [43] (PCA + area filtering + median in the neighborhood + weighted summation kernel SVM classif)</p>	<p>★ 2004: Dell’Acqua <i>et al.</i> [35] (PCA + MP + DBFE or DAFE + neural network classif)</p> <p>★ 2005: Benediktsson <i>et al.</i> [11] (PCA + EMP + DAFE or DBFE or NWFE + neural network classif)</p>
Segmentation map	<p>★ 1976: Kettig and Landgrebe [74] (region growing segmentation + ML classif of every region)</p> <p>★ 2006: Aksoy [4] (LDA + PCA + Gabor + Bayesian classif + split-and-merge segmentation + Bayesian classif)</p>	<p>★ 2007: Linden <i>et al.</i> [88] (eCognition segmentation + SVM classif of a vector mean of every region)</p> <p>★ 2009: Huang and Zhang [65] (NMF + multiscale MS or FNEA segmentation + SVM classif)</p>	<p>★ 2008: Noyel [103] (spectrum modeling + LDA + marker selection by filtering + watershed)</p>

Table 1.1 summarizes and proposes a classification of the keystone existing methods for joint spectral/spatial classification of hyperspectral data. The methods are grouped based on the used neighborhood system and the pattern recognition method. In the following, we briefly describe the evolution of spectral-spatial classification methods for hyperspectral images.

Perhaps the most intuitive way to define a pixel neighborhood is to assume that for a given pixel, its closest neighbors belong with a high probability to the same object, *i.e.*, to the same class [117]. Figure 1.5

gives examples of the *closest fixed neighborhood* systems. Four- and eight-neighborhoods are the most frequently used in image analysis. By using this approach, the pixel in the center can be classified taking into account the information from its neighbors according to one of these systems.

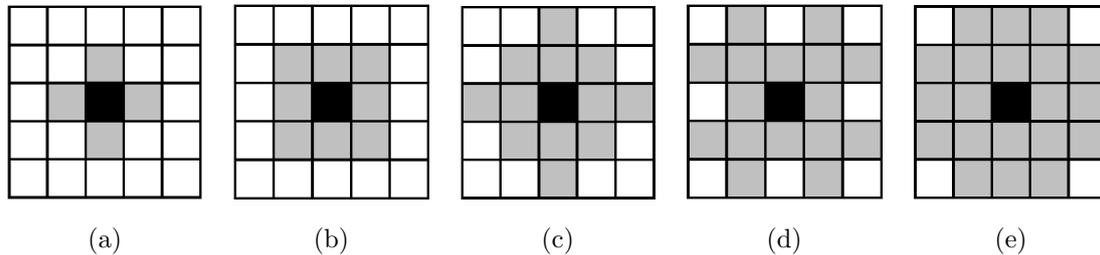


Figure 1.5: Closest fixed neighborhoods (in gray) for a black pixel: (a) four, (b) eight, (c) twelve, (d) sixteen (Chamfer neighborhood) and (e) twenty neighbors.

Following this approach, Pony *et al.* [114] and Jackson and Landgrebe [67] have proposed spectral-spatial iterative statistical classifiers for hyperspectral data based on MRF modeling. MRFs are probabilistic models that are widely used to integrate spatial context into image classification problems [39]. In the proposed method, pixelwise ML classification is firstly performed (if necessary, dimensionality reduction is previously applied). Then, the classification map is regularized using the *maximum a-posteriori* (MAP)-MRF framework. In this framework, the MAP decision rule is formulated as the minimization of an energy function consisting of the spectral and the spatial energy terms [128]. The spectral energy term is derived from the class-conditional probability for a given pixel, while the spatial energy term is computed over the pixel neighborhood. In 2005, Farag *et al.* have further explored the MAP-MRF classification. They have proposed a similar method, where class-conditional PDFs are estimated by the Mean Field-based SVM regression algorithm. As a conclusion, MRF-based methods have proved to be a powerful tool for contextual image analysis. However, they traditionally require an iterative optimization step, which is time consuming.

Dell'Acqua *et al.* have extended the neural network fuzzy ARTMAP classifier for spectral-spatial classification of hyperspectral images [35, 51]. After applying DAFE and pixelwise classification, the output is further refined by a second classification, using the same fuzzy ARTMAP method. The input vectors for the second classification represent the mapping patterns in a window around each pixel. This method has shown to be effective in a variety of situations. However, it has the same drawbacks as other neural network approaches (high computational complexity and limited generalization capability), which became less popular in the last years.

With the increasing popularity of the kernel-based SVM techniques for hyperspectral image analysis, it is interesting to incorporate the spatial context into SVM-based classifiers. Camps-Valls *et al.* have proposed a framework of composite kernels for this purpose [22]. First, for each pixel both spectral and spatial feature vectors are separately extracted. For instance, the mean and standard deviation (*per* spectral band) can be estimated over the pixel closest neighborhood and used as spatial features. The spectral and spatial feature vectors are used for computing two different kernel matrices, applying any suitable kernel function that fulfills Mercer's conditions. Finally, the two kernel matrices are combined in a unified SVM classification framework, applying one of the proposed composite kernel approaches, such as a stacked features approach, a direct summation kernel, a weighted summation kernel, a cross-information kernel and kernels for improved versatility. This methodology has shown good performances in integrating contextual information into classification and has been further exploited when considering more advanced neighborhood systems, as discussed hereafter.

Tsai *et al.* [145] and Huang and Zhang [65] have investigated the use of texture measures derived from the Gray Level Co-occurrence Matrix (GLCM) for including the spatial information in classification of hyperspectral data. In the method proposed by Tsai *et al.* [145], texture images are generated

using four measurements to describe the GLCM: Angular Second Moment, Contrast, Entropy and Homogeneity. Then, a Principal Component Analysis (PCA) is applied on the obtained texture images, and the Principal Components (PCs) are selected as features for ML classification. Huang and Zhang [65] have proposed to perform Non-negative Matrix Factorization (NMF) feature extraction first, then extract spatial information using four measurements for the GLCM (Angular Second Moment, Entropy, Homogeneity and Dissimilarity), and apply SVM classification using spatial and spectral stacked features. The experimental results were mostly not an improvement over the pixelwise ones. This may be explained by the fact that the considered remote sensing images did not contain (or contained only a few) textured regions.

As a conclusion, the advantage of the *closest fixed neighborhood* approach is its simplicity. No additional step is required to extract the neighborhoods, since they are already defined by the fixed system. Furthermore, the neighborhood includes only a small number of pixels, thus simplifying the further incorporation of spatial information into a classifier. From the analysis of previous works (see Table 1.1), it can be concluded that MRF probabilistic models are widely used for taking the closest neighborhood information into account when classifying an image. On the other side, SVM have been applied in the most recent spectral-spatial classification techniques. Therefore, it would be interesting to further exploit the combination of SVM and MRF-based techniques for context classification. Bovolo and Bruzzone [16] and Liu *et al.* [89] have investigated the integration of the SVM classification technique within an MRF framework for analysis of the Synthetic Aperture Radar (SAR) and multispectral (four-band) images. Hence, it is of interest to extend these techniques to the case of hyperspectral images.

Unfortunately, the *closest fixed neighborhoods* do not always accurately reflect information about spatial structures. For instance, they provoke assimilation of regions containing only few pixels with their larger neighboring structures and do not provide accurate spatial information at the border of regions.

Benediktsson *et al.* have proposed to use *morphological filters* as an alternative way of performing joint classification [12]. Mathematical morphology aims at analyzing spatial relations between sets of pixels, *i.e.*, extracting information about the size, shape and orientation of structures. Thus, rather than defining a crisp neighbor set for every pixel, morphological filters enable one to analyze the neighborhood of a pixel according to the structures to which it belongs.

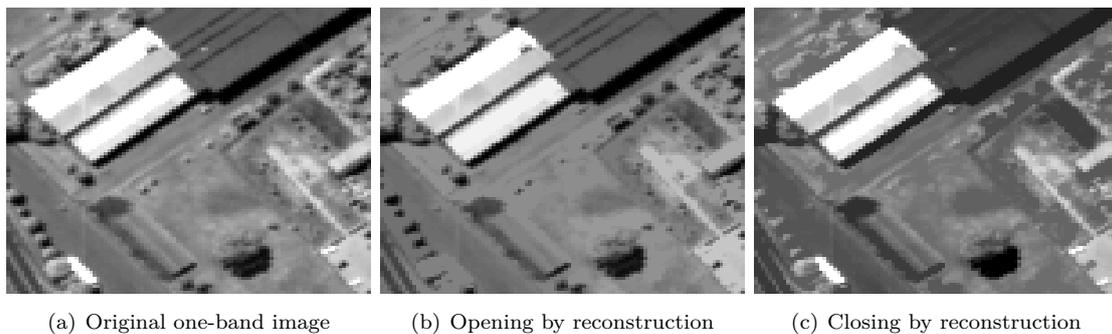


Figure 1.6: Opening and closing by reconstruction with the disk SE of radius 3 pixels.

The basic idea of mathematical morphology is to locally compare structures within the image with a reference shape called the Structuring Element (SE) [121, 126]. By using SEs of different shapes and sizes, different structures in the image scene can be analyzed.

In order to analyze spatial structures in remote sensing images, Benediktsson *et al.* have constructed a *Morphological Profile (MP)*, stemming from the granulometry principle [12, 121, 126]. An MP is composed of an Opening Profile (OP) and a Closing Profile (CP). The OP at a pixel \mathbf{x} is defined as a d -dimensional vector:

$$OP_i(\mathbf{x}) = \gamma_R^{(i)}(\mathbf{x}), \quad \forall i \in [0, d] \quad (1.15)$$

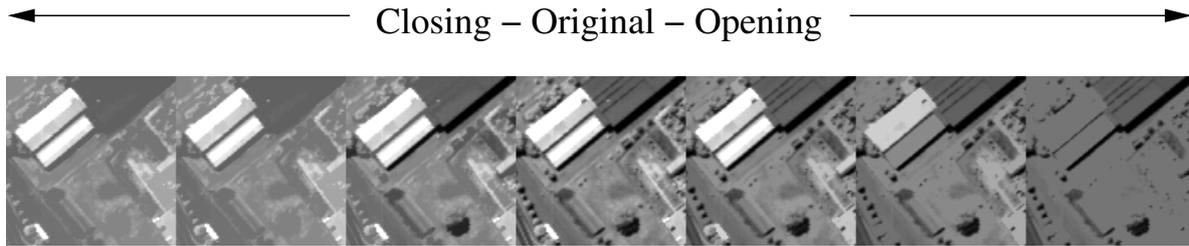


Figure 1.7: MP with the disk SE of radius $\{2,8,12\}$ pixels.

where $\gamma_R^{(i)}$ is the opening by reconstruction using an SE of size i and d is the total number of openings. The CP at a pixel \mathbf{x} is defined as a d -dimensional vector:

$$CP_i(\mathbf{x}) = \phi_R^{(i)}(\mathbf{x}), \forall i \in [0, d] \quad (1.16)$$

where $\phi_R^{(i)}$ is the closing by reconstruction using an SE of size i . Opening (resp. closing) by reconstruction is a morphological connected operator satisfying the following assertion: If a structure is brighter (resp. darker) than its immediate surrounding and cannot contain the SE, then it is totally removed, otherwise it is totally preserved. Figure 1.6 illustrates the opening and closing by reconstruction of a one-band urban image, with a disk SE of radius 3 pixels. It can be observed that small locally bright (resp. dark) objects have disappeared after the opening (resp. closing). To determine spatial structures of different shapes and sizes, it is necessary to use a range of different SEs. This assumption leads to the MP:

$$MP(\mathbf{x}) = \{CP_d(\mathbf{x}), \dots, f(\mathbf{x}), \dots, OP_d(\mathbf{x})\} \quad (1.17)$$

An example of the MP with the disk SE of different radiuses is shown in Figure 1.7. The MP was originally proposed for grayscale images. Dell'Acqua *et al.* [35] have applied the MP approach for classification of hyperspectral images. For this purpose, the first PC of multi-band data is computed (yielding an optimal one-band representation of data in the mean squared sense) [117] and the resulting grayscale image is used for the construction of an MP. The resulting MP is classified by the fuzzy ARTMAP method. The classification results have been compared for the classification of an MP with no feature extraction, a reduced MP after DBFE and after DAFE. The MP after DBFE classification has yielded the best accuracies. However, this method has not outperformed the results of the spectral fuzzy ARTMAP classification or spatial fuzzy ARTMAP using the closest fixed neighborhoods.

As a matter of fact, by using the first PC for building the MP and further classification, only a single band image has been considered, which has evidently contained a very reduced spectral information compared to the original hyperspectral data cube. To deal with this problem, Benediktsson *et al.* [11] have proposed to extract several images that contain parts of the spectral information, and then construct an MP on each of the individual images. This approach is called **Extended Morphological Profile (EMP)**. Benediktsson *et al.* [11] have built MPs on the first PCs of the hyperspectral data, and used the EMP as a single stacked vector for neural network classification (after an optional feature extraction). Later, Fauvel [43] has further investigated the EMP approach. He has proposed to perform feature extraction on the original hyperspectral data and on the EMP, and then apply SVM classification using spectral and spatial stacked features [22]. This spectral/spatial combination method enables making more use of the spectral information when compared to previous EMP-based approaches.

Although morphological opening and closing filters have shown good performances in extracting spatial structures, they cannot provide complete spatial information for an image scene, because they only act on the *extrema* of the image. Moreover, the same type of objects may appear brighter than their neighborhood in some parts of the image but darker in others. In order to analyze all the structures in the image, independently of their gray-level values, Soille has proposed to use self-complementary filters, substituting the image extrema paradigm with the more general concept of flat zones [127]. The

self-complementary area filter removes small structures from the image based on an area criterion, and yields a map of flat connected zones. Fauvel [43] has exploited these filters in a new spectral-spatial classification method for hyperspectral images. Each connected zone of the self-complementary area filtering output (the filter must be applied on a one-band image, for instance on the first PC, and the *minimum* size of a connected zone must be set by the user) defines an adaptive neighborhood for all the pixels within this zone. Then, the vector median [7] in each neighborhood is computed, and is used as the spatial feature vector for all the pixels within the considered neighborhood. Finally, SVM classification is performed, applying the weighted summation kernel for combining spectral and spatial information. The reported classification results are comparable to the ones obtained by the EMP-based technique [43].

No literature has been found on the Bayesian-based context classification using neighborhoods extracted by morphological operators. The only recent work [28] proposes to classify multispectral images by decomposing each object into bulge, disc and stellar bar structures and then reconstructing them using multichannel observations with a Bayesian model and Monte Carlo Markov Chain sampling. The authors also give perspectives for analyzing high-dimensional ultraspectral data (containing more than a thousand spectral channels) using similar approaches.

As a conclusion, morphological and area filters enable much more accurate definition of pixel neighborhoods and spatial structures in the image scene than the previously described fixed closed neighborhoods. They are especially suitable for analyzing images of urban areas containing (mostly man-made) structures with a well-defined shape. Examples of these objects are roads, buildings and trees, that can be characterized by line, rectangular and circle shapes, respectively. However, morphological operators are originally defined for one-band images. Although their extension to the case of hyperspectral data have been proposed, mostly by extracting one or a few representative grayscale images and processing these images independently, mathematical morphology for hyperspectral images needs to be further developed and defined. It can be also noted that the described methods typically add to the spectral feature vector a new extracted spatial feature vector, and classify each pixel using both vectors. This often leads to the increase of data dimensionality, additional feature extraction step and hence, implies high computational complexity of the existing techniques. Thus, it is of interest to develop new approaches for combining the spatial and the spectral information into a classifier, which would overcome the mentioned drawbacks. Furthermore, spatial neighborhoods extracted by means of morphological or area filtering are no more fixed, but they are *scale dependent*. Sizes of SEs in the case of morphological filtering and area criterion in the case of area filtering must be chosen, and they limit the sizes of spatial structures under analysis. Morphological filters cannot extract information about objects larger than the largest SE, while area filters remove all the elements with the size smaller than the chosen *minimum* area. This leads to imprecisions in the extracted spatial information.

How can these imprecisions be further reduced? Let us recall how the ECHO classifier [74], one of the first and the most popular spectral-spatial techniques, extracts spatial neighborhoods. The ECHO algorithm partitions an image into statistically homogeneous regions, *i.e.*, performs an *image segmentation*.

Definition 1.2 (Segmentation) *Segmentation is defined as an exhaustive partitioning of the input image into non-overlapping regions, each of which is considered to be homogeneous with respect to some criterion of interest (homogeneity criterion, e.g., intensity or texture) [69].*

Hence, each region in the segmentation map defines a spatial neighborhood for all the pixels within this region. This approach extracts large neighborhoods for large homogeneous regions, while not missing one or a few-pixels regions. Thus, if an accurate map of regions, corresponding to the spatial structures in the image scene, could be found, the resulting segmentation map would provide complete and precise spatial information. However, automatic segmentation of hyperspectral images is also a challenging problem.

The ECHO segmentation strategy, based on a conjunctive partitioning algorithm, requires two thresholds of statistical similarity to be set. The accuracy of segmentation results strongly depends on the selected thresholds. Furthermore, the ECHO method is based on statistical computations and involves

estimation of covariance matrices. Therefore, this approach may be not well adapted for hyperspectral data.

Aksoy [4] has proposed another scheme for spectral-spatial classification of hyperspectral data involving a segmentation step. First, spectral and textural features are extracted by means of LDA, PCA and Gabor filtering [91]. These features are quantized and used to train Bayesian classifiers with discrete non-parametric density models. Then, region segmentation is performed using an iterative split-and-merge algorithm. This approach first applies an area filtering on a pixelwise classification map. The resulting segmentation map is further processed using erosion and dilation morphological operators to automatically divide large regions into more compact sub-regions. Then, the resulting regions are modeled using the statistical summaries of their spectral and spatial (Gabor bands and ten shape features of each region, such as area, eccentricity, solidity, extent, and others) properties, and are used with Bayesian classifiers to compute the final classification map. Excellent classification accuracies prove the interest to perform image segmentation starting from supervised classification results. However, although Aksoy proposes to define spatial structures in the image using the segmentation framework, the described segmentation technique is based on morphological filtering and is similar to the method of Fauvel [43] involving self-complementary area filtering. Therefore, this approach also favors removing structures with areas smaller than the selected threshold. Furthermore, the proposed classification framework is composed of many techniques, including three feature extraction techniques, two-step iterative segmentation, quantization and classification applied twice, and hence, it is computationally demanding.

Linden *et al.* [88] have applied an SVM method for classification of every region in a hyperspectral image. First, a region growing segmentation is performed, using the eCognition software (where different thresholds lead to different segmentation maps) [34]. Then, a vector mean for every region is computed such that the value in each band represents the average spectral information of the pixels in this region in the respective band. Afterwards, the regions are classified by an SVM classifier, which efficiently handles high-dimensional data. The obtained experimental results were generally not an improvement over those obtained by the pixelwise classification. The possible reason may lie in the fact that the use of a vector mean for classifying every region significantly reduces the amount of spectral information presented in the region. This may yield imprecisions of classification. Therefore, it is of interest to explore other methods for combining the extracted spatial information with spectral information for an accurate classification.

Therefore, further developments of context classification methods for hyperspectral images, using the spatial information derived from region segmentation results, are required. Recent studies of Noyel [103] and Huang and Zhang [65] in spectral-spatial classification of hyperspectral data, conducted at the same time with this PhD thesis work, prove a great interest to this topic.

Noyel [103] has investigated morphological segmentation and classification of hyperspectral images. First, image dimensionality is reduced by spectrum modeling. Then, LDA pixelwise classification is performed. The obtained classification map is further filtered class by class, using morphological opening, in order to select large spatial regions as seeds, or *markers*, for watershed region growing segmentation [56]. The author has also proposed to use random balls (connected sets of pixels of randomly selected sizes) extracted from these large regions as markers. Since the final objective was to detect specific structures, such as glue occlusions and cancerous growth, specific mean parameters (derived from spectrum modeling) of every region were analyzed, leading to the final map of extracted objects and background.

Huang and Zhang [65] have recently published a comparative study of several spatial approaches for hyperspectral image classification. The authors have proposed a new multiscale segmentation-based technique. First, NMF feature extraction is applied [83]. Then, image segmentation is performed, using either Mean Shift (MS) or Fractal Net Evolution Approach (FNEA). The MS segmentation procedure is a non-parametric technique, which detects modes of a density function and partitions an image into clusters assuming each mode to be a center of the corresponding cluster [29]. The FNEA is a region merging technique, which initially assumes each pixel to represent one region and then iteratively merges adjacent regions satisfying a homogeneity criterion, until convergence is reached [18]. Since performances of both segmentation techniques depend on the selected parameters (a region homogeneity criterion for the FNEA and a spatial bandwidth for the MS approach, respectively), Huang and Zhang have proposed to

use multiscale segmentation results, obtained by running algorithms with d different values of parameters. The resulting spatial feature set for a given pixel contains d mean vectors of the regions containing the considered pixel at different segmentation scales. Finally, SVM classifier is used to interpret the obtained feature set. The proposed multiscale approach has been compared with spectral-spatial SVM classification using Derivative MPs (DMPs), the method designed by Fauvel and tested for panchromatic images in [43]. Excellent classification results are reported, where DMP-based technique slightly outperforms the multiscale MS approach.

Thus, both multiscale approaches (Huang and Zhang [65]) and a marker-based watershed segmentation (Noyel [103]) are designed for extracting information about spatial structures in the image, while avoiding strong dependence of the results on the selected threshold parameters. Although multiscale approaches provide segmentation maps at different scales and thus yield detailed spatial information, they are computationally and space demanding. Tilton has proposed an Hierarchical SEGmentation (HSEG) approach for hyperspectral images based on region growing and spectral clustering, which also results in a set of several image segmentations at different levels of detail [143, 142]. In order to reduce computational demands, a Recursive approximation of HSEG (RHSEG) and its efficient parallel implementation have been developed [142]. On the other side, region growing segmentation based on automatically derived markers is a promising technique for extracting spatial structures, since it allows one to mitigate dependence of performances on the selected thresholds and to avoid high computational loads.

As a conclusion, context classification approaches have shown their advantages over pixelwise techniques, providing more accurate and homogeneous classification maps. However, the existing methods suffer from scale limitations and/or algorithmic and computational complexity. Therefore, it is of great interest to further develop efficient and fast spectral-spatial methods and algorithms for analysis and classification of hyperspectral images.

1.3 Objectives

The **main objective** of this thesis is to develop new methods and algorithms for spectral-spatial classification of hyperspectral data. The developed approaches should be efficient both in terms of classification accuracies and computational complexity. Based on the discussion from the previous section, we have defined the following specific **objectives**:

1. To combine SVM and MRF-based approaches for context classification of hyperspectral images.
2. To develop classification methods using the spatial information derived from region segmentation results. For this purpose, we aim
 - (a) To explore unsupervised segmentation techniques and extend them to the case of hyperspectral images, if necessary.
 - (b) To develop rules for combining the extracted spatial information with spectral information into a classifier.
 - (c) To design classification methods using region growing segmentation based on automatically derived markers for extracting spatial structures.
3. To try avoiding usually time-consuming feature extraction/selection procedure when developing new classification approaches. For this purpose, the developed methods must be able to handle high-dimensional data. In particular, SVM classifier is a good choice of a pattern recognition approach to be applied.
4. In order to further reduce computational loads, to explore possibilities of high-performance parallel computing on commodity processors.

1.4 Main Contributions

The main contributions of this thesis are summarized in Figure 1.8, which depicts the proposed classification of the new methods and algorithms for spectral-spatial classification of hyperspectral data.

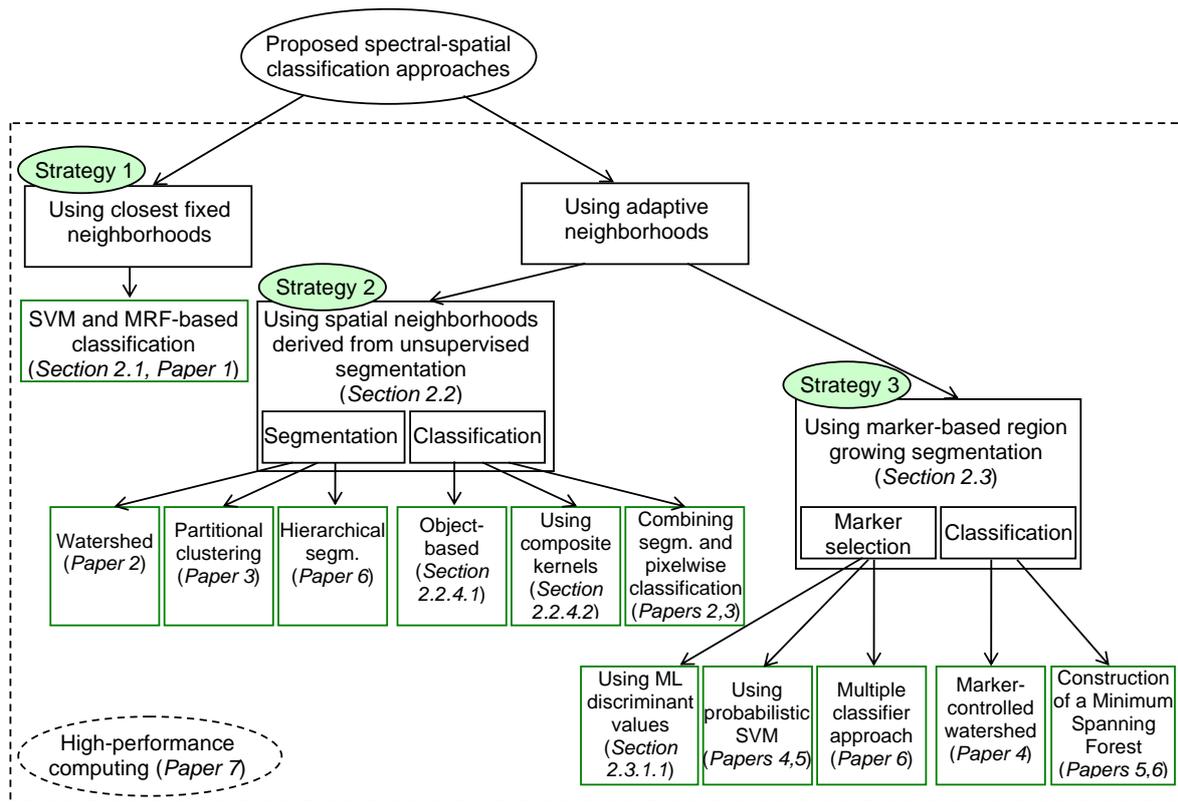


Figure 1.8: Proposed spectral-spatial classification methods for hyperspectral data (“segm.” means segmentation).

In order to achieve the objectives defined in the previous section, we have proposed and developed *three general strategies* for hyperspectral data classification:

1) The first strategy (addressed in Section 2.1 and detailed in Paper 1) explores spectral-spatial classification using closest fixed neighborhoods and is based on the integration of the SVM technique within an MRF framework. Our proposed method consists in performing a probabilistic SVM classification, followed by MRF-based spatial regularization. An important novelty consists in integrating the “fuzzy no-edge/edge” function into the spatial energy term involved in MRFs, aiming at preserving edges while performing regularization.

2) The second strategy (addressed in Section 2.2 and in Papers 2, 3, and 6) proposes to use adaptive spatial neighborhoods derived from segmentation results. Different segmentation techniques (watershed, partitional clustering, and the RHSEG methods) are investigated and extended to the case of hyperspectral images. Then, approaches for combining the extracted spatial regions with spectral information in a classifier are proposed and developed.

3) In the third strategy (described in Section 2.3 and in Papers 4-6), we have concentrated on techniques to reduce oversegmentation in an image, which is achieved by automatically “marking” the meaningful spatial structures before performing a marker-controlled segmentation. An important contribution consists in analyzing probabilistic classification results for selecting the most reliably classified pixels as markers of spatial regions. Several marker selection approaches are proposed, using either individual classifiers, or a multiple classifier system. Then, different approaches for marker-controlled region growing are developed, using either watershed or Minimum Spanning Forest methods and resulting in both

segmentation and context classification maps.

Finally, we have explored (in Section 2.5 and in Paper 7) possibilities of high-performance parallel computing using commodity processors for reducing computational loads. The developed techniques have proven to be efficient both in terms of classification accuracies and computational complexity, and thus show great potential for various image analysis scenarios.

1.5 Thesis Outline

This thesis is composed as a collection of publications, with the extended research summary. In the first part, Chapter 2 presents a summary of new methods and algorithms for spectral-spatial classification of hyperspectral images developed in the frame of this thesis. The chapter terminates with an experimental comparison of the proposed methods. Chapter 3 concludes this thesis with a discussion of main contributions and offers some suggestions for future work. Finally, this part closes with an extensive summary in French and a bibliography.

Part II is composed of seven publications presenting the main findings of this thesis:

Paper 1. Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, “SVM and MRF-based method for accurate classification of hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, 2010, DOI 10.1109/LGRS.2010.2047711.

This paper presents a novel SVM and MRF-based method for spectral-spatial classification of hyperspectral data (Figure 1.8, Strategy 1). The method consists in performing a probabilistic SVM pixelwise classification, followed by MRF-based regularization for incorporating spatial and edge information into classification.

Paper 2. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, “Segmentation and classification of hyperspectral images using watershed transformation,” *Pattern Recognition*, vol. 43, no. 7, pp. 2367-2379, July 2010.

In this paper, we propose to extend the watershed segmentation algorithm for hyperspectral images, in order to define information about spatial structures (Figure 1.8, Strategy 2). The extracted spatial regions are further used as adaptive neighborhoods for context classification. The proposed classification scheme is based on a pixelwise SVM classification, followed by majority voting within the watershed regions.

Paper 3. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, “Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2973-2987, Aug. 2009.

This paper further explores the use of adaptive spatial neighborhoods derived from segmentation results for hyperspectral image classification (Figure 1.8, Strategy 2). The proposed method combines the results of a pixelwise SVM classification and a segmentation map obtained by partitional clustering. This is achieved by performing a majority voting on the pixelwise spectral classification using adaptive neighborhoods defined by the segmentation map. The use of both the ISODATA and the Gaussian mixture resolving techniques for image segmentation are investigated. The remaining noise in the classification map is further reduced by a fixed-window-based postfiltering.

Paper 4. Y. Tarabalka, J. Chanussot, J. A. Benediktsson, “Classification based marker selection for watershed transform of hyperspectral images,” in Proc. of *IGARSS’09*, Cape Town, South Africa, 2009, pp. III-105 - III-108.

In this paper, we propose to analyze probabilistic classification results for selecting the most reliably classified pixels as markers for watershed segmentation (Figure 1.8, Strategy 3). Each marker defined from classification results is associated with a class label. By assigning the class label of each marker to all the pixels within the region grown from this marker, a spectral-spatial classification map is obtained.

Paper 5. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, “Segmentation and classification of hyperspectral images using Minimum Spanning Forest grown from automatically selected markers,” *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, DOI 10.1109/TSMCB.2009.2037132.

This paper presents a new classification method, which is based on the construction of a Minimum Spanning Forest (MSF) from region markers (Figure 1.8, Strategy 3). Markers are defined automatically from probabilistic pixelwise classification results. Each tree in the MSF forms a region in the segmentation map. A classification map is obtained by assigning a class of each marker to all the pixels within the region grown from this marker.

Paper 6. Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, “Multiple spectral-spatial classification approach for hyperspectral data,” *IEEE Trans. on Geoscience and Remote Sensing*, *under review*.

In this paper, a new multiple classifier approach for spectral-spatial classification of hyperspectral images is proposed (Figure 1.8, Strategy 3). Several classifiers are used independently to classify an image. For every pixel, if all the classifiers have assigned this pixel to the same class, the pixel is kept as a marker of the spatial region, with the corresponding class label. We propose to use spectral-spatial classifiers at the preliminary step of the marker selection procedure, each of them combining the results of a pixelwise classification and a segmentation map. Different segmentation methods based on dissimilar principles lead to different classification results. Furthermore, an MSF is built, where each tree is rooted on a classification-driven marker and forms a region in the classification map.

Paper 7. Y. Tarabalka, T. V. Haavardsholm, I. Kasen, and T. Skauli, “Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing,” *Journal of Real-Time Image Processing*, vol. 4, no. 3, pp. 287-300, Aug. 2009.

This paper investigates the use of Graphics Processing Units (GPUs) for real-time processing of large data volumes recorded by a hyperspectral imager. In particular, the paper studies a hyperspectral anomaly detection algorithm based on normal mixture modeling of the background spectral distribution. The algorithm is analyzed with respect to both complexity and potential for parallelization, and its computationally demanding parts are implemented on an Nvidia GeForce 8800 GPU using the Compute Unified Device Architecture programming interface.

Chapter 2

Proposed Methods

THIS chapter presents a summary of the new methods and algorithms developed in the frame of this thesis and thoroughly discussed in Part II. First, the proposed approaches for spectral-spatial classification of hyperspectral data are described. Then, the results of an experimental comparison of the proposed methods are reported. Finally, possibilities of speeding up computational processes using high-performance parallel computing on commodity processors are discussed.

2.1 Context Classification Using Closest Fixed Neighborhoods, SVM and MRFs

We first explore context classification using closest fixed neighborhoods to analyze spatial dependencies between pixels. As was discussed in the previous chapter, SVM and MRF probabilistic models are two powerful tools for high-dimensional data classification and for contextual image analysis, respectively. Bovolo and Bruzzone [16] and Liu *et al.* [89] have developed MRF and SVM-based methods for classification of SAR and multispectral (four-band) images, respectively. The authors use SVM to estimate class conditional PDFs and MRFs to estimate context-based class priors for the final MAP decision rule. We propose to extend this approach to the case of hyperspectral data.

In the following, we present a novel SVM and MRF-based method (*SVMMRF*) for spectral-spatial classification of hyperspectral images (described in Paper 1). In the first step of the proposed method, a probabilistic SVM pixelwise classification is applied [152, 25]. In the second step, spatial contextual information is used for refining the classification results obtained at the first step. This is achieved by means of the MRF regularization. An important difference from previously proposed methods [42, 16, 89] consists in defining and integrating the “fuzzy no-edge/edge” function into the spatial energy function involved in MRF, aiming at preserving edges while performing spatial regularization.

The flowchart of the proposed *SVMMRF* classification method is shown in Figure 2.1. At the input a B -band hyperspectral image is available, which can be represented as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. We recall that classification consists in assigning each pixel to one of the K classes of interest $\{\omega_1, \omega_2, \dots, \omega_K\}$.

2.1.1 Probabilistic SVM classification

The first step of the proposed procedure consists in performing a probabilistic SVM pixelwise classification of a hyperspectral image [110, 148]. Other classifiers could be used. However, SVM are extremely well suited to classify hyperspectral data [20]. We have described the SVM classification technique in the previous chapter (see Section 1.2), and we refer the reader to [20, 148] for a more detailed analysis of the SVM. The outputs of this step are the following:

1. a classification map, where each pixel has a unique class label;

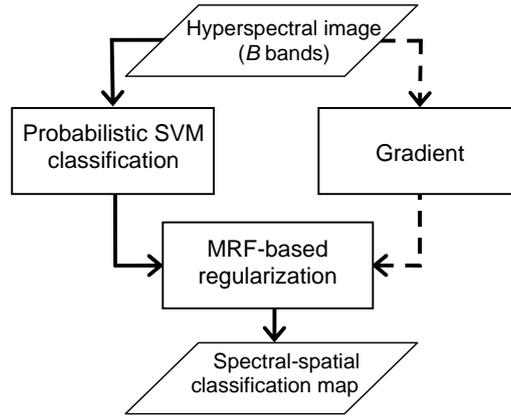


Figure 2.1: Flowchart of the proposed *SVMMRF* classification scheme.

2. probability estimates for each pixel to belong to each class of interest.

The standard SVM do not provide probability estimates for the individual classes. In order to get these estimates, pairwise coupling of binary probabilistic estimates can be applied [25, 152]. Two techniques for computing probability estimates for multi-class classification by pairwise coupling are described in [152]. We propose to use one of these methods which is implemented in the LIBSVM library [25]. The objective is to estimate, for each pixel \mathbf{x} , the probabilities to belong to each class of interest:

$$\mathbf{p} = \{p_k = p(y = k|\mathbf{x}), k = 1, \dots, K\}. \quad (2.1)$$

For this purpose, first multiclass “one versus one” SVM classification with the Gaussian RBF kernel is applied. Pairwise class probabilities $r_{ij} \approx p(y = i|y = i \text{ or } j, \mathbf{x})$ are estimated, using an improved implementation [87] of [110]:

$$r_{ij} \approx \frac{1}{1 + e^{A\hat{f} + B}}, \quad (2.2)$$

where A and B are estimated by minimizing the negative log-likelihood function using known training data and decision values \hat{f} . Furthermore, the probabilities in (2.1) are computed by solving the following optimization problem:

$$\min_{\mathbf{p}} \sum_{i=1}^K \sum_{j:j \neq i} (r_{ji}p_i - r_{ij}p_j)^2 \text{ subject to } \sum_{i=1}^K p_i = 1, p_i \geq 0, \forall i. \quad (2.3)$$

This problem has a unique solution, which can be found by solving a simple linear system, as described in [152].

2.1.2 Computation of the gradient

Independently of the previous step, a one-band gradient of the hyperspectral image is computed. Gradient defines transitions between regions, so that it has high values on the borders between objects and local *minima* in the homogeneous regions. We further use the gradient for defining the fuzzy no-edge/edge function.

Approaches for computing a one-band gradient from the B -band image are analyzed in Section 2.2.1 [104, 140]. We propose to use the following approach in the *SVMMRF* method: First, we compute horizontal, vertical and two diagonal gradients (corresponding to the directions 0° , 90° , 45° and 135° , respectively), using Sobel masks [56], where each of the gradients is computed as the sum of the gradients of every spectral channel. The resulting one-band gradient $\nabla(\mathbf{X}) = \{\rho_j \in \mathbb{R}, j = 1, 2, \dots, n\}$ is found as the average of the four obtained directional gradients.

2.1.3 MRF-based regularization

In the final step, the regularization of the SVM classification map is performed, using the MAP-MRF framework. This framework is based on the interpixel class dependency assumption, meaning that a pixel belonging to a class ω_i is likely to have neighboring pixels belonging to the same class. For instance, an eight-neighborhood can be assumed (let \mathbf{N}_i be the set of neighbors for a given pixel \mathbf{x}_i).

We adopt the Metropolis algorithm, based on stochastic relaxation and annealing, for computing the MAP estimate of the true classification map given the initial pixelwise classification map [52, 96]. The considered method is based on the Bayesian approach, and aims at minimizing the global energy in the image, by iterative minimization of local energies associated with randomly chosen image sites, *i.e.*, pixels.

Let $\mathbf{L} = \{L_j, j = 1, 2, \dots, n\}$ be a generic set of information class labels for the image \mathbf{X} . We propose to compute the local energy of a given site associated with a pixel \mathbf{x}_i as

$$U(\mathbf{x}_i) = U_{spectral}(\mathbf{x}_i) + U_{spatial}(\mathbf{x}_i), \quad (2.4)$$

where $U_{spectral}(\mathbf{x}_i)$ is the spectral energy function from the observed data and $U_{spatial}(\mathbf{x}_i)$ is the spatial energy term computed over the local neighborhood \mathbf{N}_i . We define the spectral energy term as

$$U_{spectral}(\mathbf{x}_i) = -\ln\{P(\mathbf{x}_i|L_i)\}, \quad (2.5)$$

where $P(\mathbf{x}_i|L_i)$ is estimated by pairwise coupling of probability estimates from “one *versus* one” SVM outputs [89, 152]. For the spatial energy term, two different expressions are proposed. We first consider the standard spatial energy expression [16], which is computed as

$$U_{spatial}^{NE}(\mathbf{x}_i) = \sum_{\mathbf{x}_j \in \mathbf{N}_i} \beta(1 - \delta(L_i, L_j)), \quad (2.6)$$

where $\delta(\cdot, \cdot)$ is the Kronecker delta function ($\delta(a, b) = 1$ if $a = b$ and $\delta(a, b) = 0$ otherwise) and β is a parameter that controls the importance of the spatial *versus* the spectral energy terms. The superscript “*NE*” means that no edge information is taken into account. The term $U_{spatial}^{NE}(\mathbf{x}_i)$ is proportional to the number of the neighboring pixels of \mathbf{x}_i assigned to one of the classes different from L_i . This spatial energy term is especially suitable for images with large spatial structures. However, if a small, one-pixel object is present in the image, this model will favor assigning this pixel to the class of the surrounding objects.

In order to mitigate this drawback of the previous spatial term, and to preserve small structures and edges in the classification map, we propose to integrate the edge information into the spatial energy function. The computation of an accurate edge map for hyperspectral images is a challenging task. For instance, it can be obtained by thresholding the gradient image $\{\rho_j \in \mathbb{R}, j = 1, 2, \dots, n\}$. For this purpose, an appropriate threshold must be chosen. Instead of computing the edge map, we propose to define the following “fuzzy no-edge/edge function”:

$$\varepsilon(\mathbf{x}_j) = 1 - \frac{\rho_j}{\alpha + \rho_j}, \quad 0 < \varepsilon(\mathbf{x}_j) \leq 1, \quad (2.7)$$

where $\alpha > 0$ is a parameter controlling the approximate edge threshold. When $\rho_j = 0$ (no edge), we have $\varepsilon(\mathbf{x}_j) = 1$. The value of $\varepsilon(\mathbf{x}_j)$ gets smaller and closer to zero as ρ_j increases. From here, the following spatial energy expression is proposed:

$$U_{spatial}^E(\mathbf{x}_i) = \sum_{\mathbf{x}_j \in \mathbf{N}_i} \beta \varepsilon(\mathbf{x}_j) (1 - \delta(L_i, L_j)). \quad (2.8)$$

The superscript “*E*” means that the edge information is taken into account. In the following, we thus refer to two different methods, namely *SVMMRF-NE* and *SVMMRF-E*, when (2.6) and (2.8) are used for computing the spatial energy, respectively.

We briefly summarize the considered Metropolis algorithm for optimizing the energy function: In each iteration, an image site (*i.e.*, a pixel \mathbf{x}_i) is randomly chosen. The local energy of the given site $U(\mathbf{x}_i)$ is computed by (2.4). Then, a new class label L_i^{new} is randomly selected for the site \mathbf{x}_i , and a new local energy $U^{new}(\mathbf{x}_i)$ is computed. If the variation of the energy $\Delta U = U^{new}(\mathbf{x}_i) - U(\mathbf{x}_i) < 0$, the new class label is assigned to \mathbf{x}_i : $L_i = L_i^{new}$. Otherwise, the new class assignment is accepted with the probability $p = \exp(-\Delta U/T)$. Here, T is a global control parameter called “temperature” [52]. The Metropolis algorithm requires choosing carefully the temperature scheme. The optimization begins at a high temperature, which is gradually lowered as the relaxation proceeds. This algorithm avoids converging to local *minima*.

2.1.4 Concluding discussion

As will be shown in Section 2.4, the proposed *SVMMRF* method yields excellent classification results for a variety of images. However, as was discussed in the previous chapter, the closest fixed neighborhoods do not accurately reflect information about the objects present in the images. Moreover, due to the use of the closest neighborhoods containing few pixels, the proposed method is efficient only if there is no large misclassified region in the initial pixelwise classification map (this assumption often holds). If such a region exists, the MRF-based method cannot reconstitute its true class label.

2.2 Spectral-spatial Classification Using Adaptive Neighborhoods

In the remaining part of the thesis, we develop contextual classification methods using flexible, or *adaptive*, spatial neighborhoods. As was previously discussed, such adaptive neighborhoods can be derived from region segmentation results. Therefore, in this chapter we explore (*unsupervised*) *segmentation* techniques, extending them to the case of hyperspectral images. Then, we develop approaches for combining the extracted spatial regions with spectral information into a context classifier.

Fu and Mui [48] identified three classes of image segmentation techniques: *Edge-based*, *region-based* and *characteristic feature thresholding* or *clustering*. Methods from the first two classes operate in the spatial domain, while those from the last class work in the spectral domain. Finally, a combination of spatial-based and spectral-based segmentation is possible. As described hereafter, we have investigated techniques from these different classes: 1) spatial-based segmentation using watershed transformation; 2) spectral-based segmentation using partitional clustering; 3) segmentation both in the spatial and the spectral domain using the Recursive Hierarchical SEGmentation (RHSEG) method.

2.2.1 Spatial-based segmentation using watershed transformation

Segmentation techniques working in the spatial domain search for groups of spatially connected pixels (*i.e.*, regions) which are similar accordingly to the defined criterion. *Edge-based* techniques search for discontinuities in the image, while *region-based* techniques search for similarities between image regions.

In previous studies, several methods for multispectral image segmentation have been investigated. Numerous works exploit region merging techniques, where neighboring image segments are merged iteratively based mostly on their spectral similarity [34, 74, 144]. For instance, the eCognition software performs multiresolution segmentation, based on a bottom-up region merging [34]. Initially, each pixel is considered as a separate region, and subsequently pairs or regions are merged, based on a homogeneity criterion, which is a combination of spectral and shape properties. The main drawback of applying region merging for image segmentation is that the homogeneity criterion or convergence criterion must be chosen (in other words, the level of details for a segmentation map must be defined). For accurate segmentation, these techniques usually produce a pyramid of segmentation maps, using a range of thresholds.

Other studies exploit mathematical morphology based segmentation approaches [3, 46, 85, 104, 109, 119, 125, 140], which mostly use granulometries or watershed transformation. The extension of morphological operators to the case of multispectral images is not straightforward, because there is no

natural means for total ordering of multivariate pixels, which is a requirement in mathematical morphology. **Watershed transformation** is a powerful mathematical morphology technique for image segmentation [15, 126]. In a recent paper, Noyel *et al.* [104] gave an overview of the literature on the watershed-based multispectral image segmentation and applied this method to hyperspectral images. Their approach is composed of a spectral classification performed in order to obtain markers and the computation of a multivariate gradient in order to get spatial information. Only visual results (*i.e.*, the obtained segmentation maps) are presented in the article. Therefore, the question of defining a watershed transformation for the case of hyperspectral images has only recently been raised in the literature and needed further investigations. We explore and propose the extension of the watershed segmentation to hyperspectral data (see Paper 2). In the following, the watershed segmentation technique is described, and then its extensions for hyperspectral images are proposed and discussed.

2.2.1.1 Watershed segmentation

Watershed transformation considers a two-dimensional one-band image as a topographic relief [15, 126]. The value h of a pixel stands for its elevation. The watershed lines divide the image into catchment basins, so that each basin is associated with one *minimum* in the image (see Figure 2.2(a)). The watershed transformation is usually applied to the gradient function of the image. We remind that the gradient has high values on the edges of image objects and low values in homogeneous regions. In this case, if the crest lines in the gradient image correspond to the borders between objects, watershed transformation partitions this image into meaningful regions.

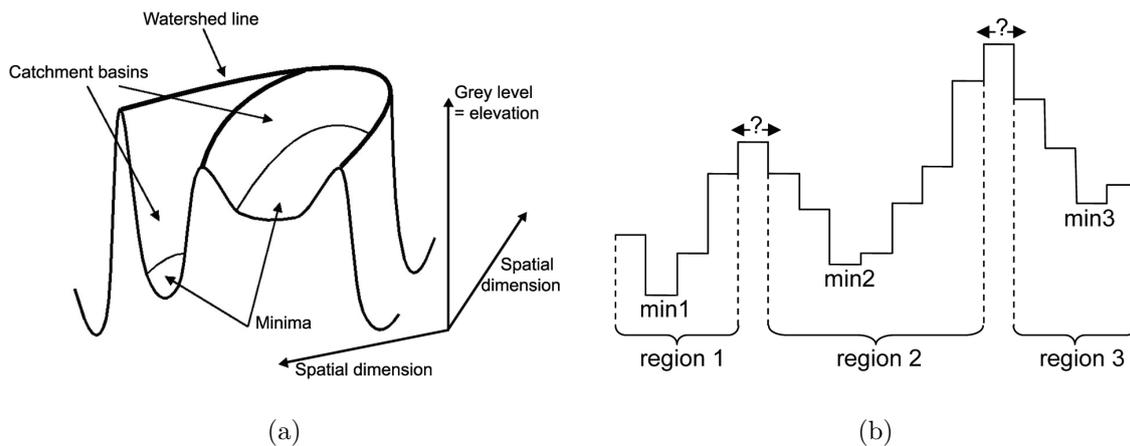


Figure 2.2: (a) Topographic representation of a one-band image. (b) Example of a watershed transformation in one dimension.

A wealth of literature describes techniques for computing the watershed transformation (see for instance [98, 106, 132, 149]). Vincent and Soille [149] have proposed an efficient algorithm using flooding simulations, which has become one of the classical algorithms to compute watersheds. The output of the watershed transform is a partition of the image composed of regions (sets of pixels connected to the same local *minimum*) and of *watershed pixels* (borders between regions). Figure 2.2(b) shows an example of watershed transformation in one dimension, where three regions, associated with the three *minima*, are defined. The two *maxima* correspond to the borders between regions and are not assigned to any region (watershed pixels).

Typically, the result of watershed segmentation on the gradient image without any additional processing is a severe oversegmentation (every single local *minimum* of the gradient leads to one region). Common ways to reduce the number of local *minima* are to filter the original image or the gradient function (*e.g.*, area filtering) or to use markers [99]. We explore marker-based watershed segmentation

in Section 2.3. The oversegmentation effect can be also corrected using some post-processing, such as merging of similar neighboring regions. In the next section, we discuss and extend different approaches for segmentation of hyperspectral images by a watershed transformation.

2.2.1.2 Watershed segmentation of hyperspectral images

As previously mentioned, the watershed transformation requires as input a one-band image. Our objective is to apply the watershed to a B -band hyperspectral image $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. We also denote the image of every spectral band as $X_\lambda, \lambda = 1, 2, \dots, B$. Different strategies of computing watershed are summarized in Figure 2.3.

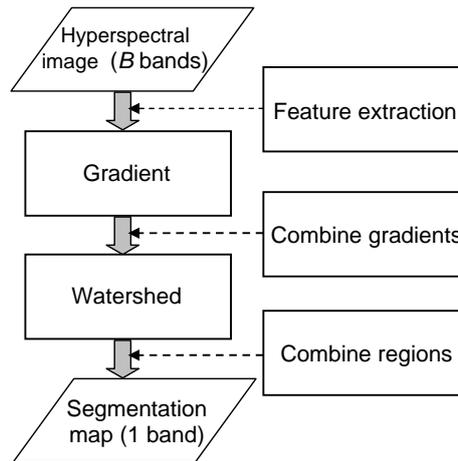


Figure 2.3: Flowchart which shows strategies of applying watershed to a hyperspectral image

Before computing a gradient, *feature extraction* on the original image can be performed, applying one of the transformations such as Principal Component Analysis (PCA) [61, 117], Independent Component Analysis (ICA) [101, 150] and Maximum Noise Fraction (MNF) [58]. The aim of this step is to obtain either a one-band image or a multi-band image which would contain enough information to distinguish between spatial structures in the image. If a one-band image with a good distinguishing capability between structures can be obtained, the algorithm for computing a gradient and watershed is straightforward. For instance, for a one-band image Y , a basic morphological gradient can be applied.

Definition 2.1 (Basic morphological gradient) *A basic morphological gradient of an image Y is defined as the arithmetic difference between the dilation and the erosion of Y by the structuring element E (definitions of the dilation and the erosion can be found in [126]):*

$$\rho_E(Y) = \delta_E(Y) - \varepsilon_E(Y). \quad (2.9)$$

If, however, at the input of the *gradient* step we still have a multi-band image, we can proceed in different ways, that can be grouped into the following three categories, discussed hereafter:

- to compute a vector gradient;
- to compute a multidimensional gradient;
- to combine watershed segmentation maps *a posteriori*.

1) Computation of a vector gradient

Vector gradients are based on the distance between pixel vectors, and produce from the B -band image a one-band gradient [41, 104]. Noyel *et al.* [104] proposed to use a metric-based gradient for hyperspectral images, which is described as follows:

Definition 2.2 (Metric-based gradient) For each pixel vector \mathbf{x}_p , let $\psi = [\mathbf{x}_p^1, \mathbf{x}_p^2, \dots, \mathbf{x}_p^e]$ be a set of e vectors in the neighborhood of \mathbf{x}_p (set ψ does not contain \mathbf{x}_p). The metric-based gradient is defined as a difference between the supremum and the infimum of the defined distances between \mathbf{x}_p and vectors from the set ψ :

$$\nabla_{\psi,d}^{MB}(\mathbf{x}_p) = \sup_{i \in \psi} \{d(\mathbf{x}_p, \mathbf{x}_p^i)\} - \inf_{j \in \psi} \{d(\mathbf{x}_p, \mathbf{x}_p^j)\}. \quad (2.10)$$

Various distances can be used to compute gradient from (2.10) such as Euclidean, Mahalabobis, chi-squared distances [104].

Another type of the vector gradient is the *Robust Color Morphological Gradient* (RCMG). It was developed for color images by Evans and Liu [41]. We propose to use the RCMG for hyperspectral images.

Definition 2.3 (Color Morphological Gradient) For each pixel vector \mathbf{x}_p , let $\chi = [\mathbf{x}_p^1, \mathbf{x}_p^2, \dots, \mathbf{x}_p^e]$ be a set of e vectors within a structuring element E , which defines the neighborhood of the vector \mathbf{x}_p , and the set χ contains \mathbf{x}_p . The Color Morphological Gradient (CMG) is defined as

$$\nabla_{\chi,d}^{CM}(\mathbf{x}_p) = \max_{i,j \in \chi} \{d(\mathbf{x}_p^i, \mathbf{x}_p^j)\}, \quad (2.11)$$

i.e., the maximum of the distances between all pairs of vectors in the set χ .

Here, various distances can be chosen. If the Euclidean distance is used, (2.11) can be rewritten as

$$\nabla_{\chi,d}^{CM}(\mathbf{x}_p) = \max_{i,j \in \chi} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}. \quad (2.12)$$

One of the drawbacks of the CMG is that it is very sensitive to noise. To overcome the problem of outliers, the RCMG has been proposed [41]. The scheme to make the CMG robust consists in removing the two pixels that are the furthest apart and then finding the CMG of the remaining pixels. This process is repeated several times until a good estimate of the gradient is obtained.

Definition 2.4 (Robust Color Morphological Gradient) The RCMG, using the Euclidean distance, is defined as

$$\nabla_{\chi,d}^{RCM}(\mathbf{x}_p) = \max_{i,j \in [\chi - REM_r]} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}, \quad (2.13)$$

where REM_r is the set of the r vector pairs removed.

The appropriate value of the parameter r in (2.13) depends on the chosen structuring element E and the amount of noise present in the image, as discussed in [41]. When a one-band vector gradient is computed, it is used as the input of the watershed algorithm.

2) Multidimensional gradient methods

Another approach to computing a one-band gradient from the multi-band image consists in considering the B -band image as a set of B one-band images. In this case, the gradients of every spectral channels can be computed, using for instance a morphological gradient (see (2.9)). Then the obtained B gradient images $\rho_E(X_\lambda)$, $\lambda = 1, 2, \dots, B$ are combined into one image using linear or non-linear operators.

As an example of the linear operators, the weighted sum of gradients can be computed by

$$\nabla_E^+(\mathbf{X}) = \sum_{\lambda=1}^B \omega_\lambda \rho_E(X_\lambda), \quad (2.14)$$

where ω_λ denotes the weight of the gradient of the band λ . If $\omega_\lambda = 1$, $\lambda = 1, 2, \dots, B$, all the bands are supposed to have an equal importance in defining the gradient. Modifying the weight coefficients, the gradient estimation can be improved. For instance, coefficients that are inversely proportional to the estimated noise of each spectral band can be used as the weights in equation (2.14).

Examples of non-linear operators are the *supremum* and the median operators. The gradient *supremum* over morphological gradients of every band is defined as follows: for every pixel the *supremum* over all gradient images is taken as the output value of this pixel.

3) Combination of watershed segmentation maps

B standard gradients are first computed, one for each spectral band. Each gradient image is then used to compute a watershed transformation. This gives B segmentation maps that can be further combined to provide a single output segmentation map.

One of the ways to combine the B segmentation maps, in order to define relevant edges, consists in summing watershed lines. Here, for each segmentation map obtained from the gradient of band λ , a binary image W_λ of watershed lines is produced. Thus, W_λ is an image in which watershed pixels are equal to 1 and all other pixels are equal to 0. The sum of the watershed lines is computed by

$$\mathbf{W} = \sum_{\lambda=1}^B W_\lambda. \quad (2.15)$$

The obtained image \mathbf{W} can be subject to further thresholding in order to define the border pixels that were presented in most of the segmentation maps, hence ensuring a reliable edge detection. However, when summing the watershed lines, we do not have information about regions but only about edges. Furthermore, some edges can become open after thresholding. Therefore, closing of edges and image region labeling must be performed after the procedure described above.

2.2.1.3 Segmentation map without watershed pixels

When one of the watershed segmentation techniques described above is applied to a hyperspectral image, in the resulting segmentation map each pixel contains a label of the region it belongs to or a watershed pixel label (see Figure 2.2(b)). It is often desirable to produce a segmented image where each pixel belongs to some region, without border pixels between regions. In this case, each watershed pixel can be assigned to one of the regions in its neighborhood. For this purpose, we propose to compute a standard vector median [7] for every region \mathbf{S} ($\mathbf{S} = \{\mathbf{s}_j \in \mathbb{R}^B, j = 1, 2, \dots, l\}$, $\mathbf{S} \subseteq \mathbf{X}$, with l equal to the number of pixels in the region).

Definition 2.5 (Standard vector median) *A standard vector median \mathbf{s}_{VM} for a set of pixel vectors is a vector which fulfills the condition that the sum of the distances between this vector and all the other vectors in the set is minimal.*

For instance, when L_1 norm is used to compute distances, vector median is computed as

$$\mathbf{s}_{VM} = \arg \min_{\mathbf{s} \in \mathbf{S}} \left\{ \sum_{j=1}^l \|\mathbf{s} - \mathbf{s}_j\|_1 \right\}. \quad (2.16)$$

Finally, every watershed pixel is assigned to the neighboring region with the “closest” median, *i.e.*, the distance between the vector median of this region and the watershed pixel vector is minimal.

2.2.2 Spectral-based segmentation using partitional clustering

Spectral-based segmentation techniques search for similarities between image pixels and clusters of pixels, not taking into consideration the spatial location of these pixels. Lambert and Macaire [79] have grouped these techniques into two classes: *Histogram-based* and *cluster-based* methods. The *histogram-based* methods relate the modes of a spectral histogram to homogeneous regions in the image [63, 107]. With a high dimensionality, these methods become memory-consuming and produce less accurate results. The *cluster-based* segmentation techniques aim at finding distinct structures in the spectral feature space.

Clustering is therefore an exhaustive partitioning of a set of pixels from the input image into homogeneous groups of pixels.

A taxonomy and survey of clustering techniques can be found in [68]. Two principal groups of clustering methods can be distinguished: Hierarchical and partitional approaches. Hierarchical methods usually produce a dendrogram, where at the lowest level each cluster contains only one pixel (*i.e.*, each pixel forms a cluster), and with the increase of levels the most similar clusters are merged (and the number of clusters decreases). Then, the result with the desired number of clusters can be chosen. Lee and Crawford [84] have applied the hierarchical clustering approach for unsupervised classification of hyperspectral images. Hierarchical clustering is a versatile technique for image segmentation that can produce a series of segmentation results. However, its application to high-dimensional data leads to significant time and memory requirements.

The alternative approach is ***partitional clustering***. The algorithms of this approach produce a single partition of the data and have advantages when applied to large data sets, as they are computationally and memory less demanding than hierarchical clustering methods. However, the number of desired clusters must be chosen. The other problem accompanying the use of these techniques is that the clustering results depend on the initialization. We have investigated two approaches of partitional clustering for hyperspectral image segmentation, namely: (1) Iterative Self-Organizing DATA Analysis (ISODATA) [8] and (2) Expectation Maximization (EM) to resolve the Gaussian mixture [36] (see Paper 3). In the following, we give a general description of the clustering procedure, followed by a brief discussion of the considered algorithms.

The three principal stages of the clustering technique are:

1. *Feature selection/extraction*: *Feature selection* consists in identifying a subset of the original features. *Feature extraction* consists in applying one or more transformations of the input features to produce new salient features. As a pixel vector from hyperspectral image contains hundreds of spectral values, feature extraction/selection is often a required step.
2. *Similarity measure*: Clustering aims at grouping pixels, so that pixels belonging to the same cluster are spectrally similar. To quantify this relationship, a similarity measure must be chosen. Proximity between pixels is usually measured by a distance function defined on pairs of spectral values. A simple distance measure like the Euclidean distance is often used to measure similarity between vectors. Description of various distance measures can be found in [37, 68, 104].
3. *Grouping*: In this step, pixels are grouped into clusters. Partitional clustering algorithms identify the partition that optimizes a clustering criterion.

Both ISODATA and EM are iterative optimization techniques. Thus, on each iteration i a partition $\mathbf{Q}_1^i, \mathbf{Q}_2^i, \dots, \mathbf{Q}_C^i$ of the set $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$ into C clusters is computed, so that $\mathbf{Q}_c^i = \{\mathbf{x}_{j,c}^i \in \mathbb{R}^B, j = 1, 2, \dots, m_c^i\}$ contains the pixels belonging to the component c on the iteration i , and m_c^i is the number of pixels in \mathbf{Q}_c^i .

2.2.2.1 ISODATA algorithm

The simplest and most frequently used criterion in partitional clustering is the squared error criterion, which is the most suitable in the case of isolated and compact clusters. The squared error for a clustering Υ of a set \mathbf{X} into C clusters is defined as:

$$e^2(\mathbf{X}, \Upsilon) = \sum_{c=1}^C \sum_{j=1}^{m_c} \|\mathbf{x}_{j,c} - \boldsymbol{\mu}_c\|^2 \quad (2.17)$$

where $\boldsymbol{\mu}_c$ is the centroid of the cluster c .

ISODATA clustering is a well-known algorithm introduced by Ball and Hall [8] which uses the squared error criterion. It starts with a random initial partition of the pixel vectors into candidate clusters and then reassigns these vectors to clusters in such a way that the squared error (2.17) is reduced at each iteration, until a convergence criterion is achieved. The algorithm permits splitting, merging and

deleting of clusters, in order to produce more accurate results and to mitigate dependence of results on the initialization.

The ISODATA algorithm is implemented in the ENVI software [1], where its application for hyperspectral images is straightforward. A vector of spectral values can be used as a feature vector for every pixel. We propose to define a *minimum* number of clusters C_{min} equal to the number of information classes in the reference map, and choose a *maximum* number of clusters C_{max} superior to this value.

2.2.2.2 EM algorithm

While ISODATA is a deterministic clustering approach, the EM algorithm belongs to the group of statistical algorithms. The underlying assumption for the mixture resolving approach to cluster analysis (that includes the EM algorithm) is that the pixel vectors are drawn from one or several distributions. The objective is to identify the parameters of each distribution. Most often the individual components of the mixture density are assumed to be Gaussian. In this case, the parameters of a Gaussian mixture model have to be estimated. The EM algorithm was proposed by Dempster *et al.* [36] to obtain iteratively a maximum likelihood estimate of the parameters of component densities from the pixel vectors.

In order to cluster a hyperspectral image by the EM technique, we assume that pixels belonging to the same cluster are drawn from a multivariate Gaussian probability distribution. Each image pixel can be statistically modeled by the following PDF:

$$p(\mathbf{x}) = \sum_{c=1}^C \omega_c \phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \quad (2.18)$$

where $\omega_c \in [0, 1]$ is the mixing proportion (weight) of cluster c with $\sum_{c=1}^C \omega_c = 1$, and $\phi(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate Gaussian density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$\phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) = \frac{1}{(2\pi)^{B/2}} \frac{1}{|\boldsymbol{\Sigma}_c|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c^{-1} (\mathbf{x} - \boldsymbol{\mu}_c)\right\}. \quad (2.19)$$

The parameters of the distributions $\boldsymbol{\psi} = \{C, \omega_c, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c; c = 1, 2, \dots, C\}$ are estimated by the EM algorithm [24], as described in Paper 3 [137]. During the procedure of parameter estimation, pixels are assigned to the C clusters. Therefore, when the algorithm converges, the partitioning of the set of pixel vectors into C clusters is obtained. An upper bound on the number of clusters, which is a required input parameter, is recommended to be chosen slightly superior to the number of classes.

The total number of parameters to be estimated is $P = (B(B+1)/2 + B+1)C + 1$, where B is a dimensionality of feature vectors. If the value of B is large, P may be quite a large number. This may cause the problem of the covariance matrix singularity or inaccurate parameter estimation results. To avoid these problems, we propose to previously perform a feature reduction, using the method of Piecewise Constant Function Approximations (*PCFA*) [71]. The simple averaging of adjacent spectral bands, applied in the PCFA algorithm, makes the resulting features directly interpretable in a physical sense. Other features extraction methods can be applied, such as PCA, DBFE and best-based algorithms [77, 81, 117].

2.2.2.3 Segmentation using clustering

The partitional clustering algorithm produces an exhaustive partitioning of the set of image pixels \mathbf{X} into C clusters. Thus, each pixel has a numerical label of the cluster it belongs to. However, since no spatial information is used during the clustering procedure, pixels with the same cluster label can be connected in the image plane (thus forming a spatial region) or they can belong to disjoint regions within the spatial coordinates. In order to obtain a segmentation map (where each connected spatial region has a unique label), a connected components labeling algorithm must therefore be applied to the output image partitioning obtained by the clustering algorithm [38, 68, 123]. This algorithm allocates different labels for disjoint regions in the image plane that were placed in the same cluster.

The obtained segmentation map can be oversegmented, as reported for example in [2]. However, the final goal is not to obtain a perfect segmentation result, but rather to define adaptive neighborhoods

(groups of connected pixels belonging to the same physical object) in order to incorporate them into a spectral-spatial classifier. Therefore, oversegmentation (one region is detected as several ones) is not a crucial problem, while undersegmentation (several regions are detected as one) is not desired. As oversegmentation is preferable to undersegmentation, a four-neighborhood connectivity is preferable to use while performing the labeling of connected components.

2.2.3 Hierarchical segmentation both in the spatial and the spectral domain

We also investigate the Hierarchical SEGmentation (HSEG) algorithm for hyperspectral images, as an example of a segmentation technique working both in the spatial and the spectral domain. This technique developed by J. Tilton [142, 143, 144] is based on iterative Hierarchical Step-Wise Optimization (HSWO) region growing method [9]. Furthermore, it provides a possibility of merging spatially non-adjacent regions by spectral clustering.

The following outline of the HSEG algorithm is based on the description given in [142, 144]:

1. Initialize the segmentation by assigning for each pixel a region label. If a pre-segmentation is provided, label each pixel according to the pre-segmentation. Otherwise, label each pixel as a separate region.
2. Calculate the dissimilarity criterion value between all pairs of spatially adjacent regions.
3. Find the smallest dissimilarity criterion value $dissim_val$ and set $thresh_val$ equal to it. Then merge all pairs of spatially adjacent regions with $dissim_val = thresh_val$.
4. If a parameter $S_{wght} > 0.0$, merge all pairs of spatially non-adjacent regions with $dissim_val \leq S_{wght} \cdot thresh_val$.
5. If convergence is not achieved, go to step (2).

In order to reduce computational demands, a Recursive divide-and-conquer approximation of HSEG (RHSEG) has been developed. The NASA-Goddard RHSEG software provides an efficient implementation of the RHSEG algorithm.

When determining the most similar pair of regions, we propose to choose the standard Spectral Angle Mapper (SAM) between the region mean vectors and as the dissimilarity criterion (see equation 1.5) [142]. The optional parameter S_{wght} tunes the relative importance of spectral clustering *versus* region growing. If $S_{wght} = 0.0$, only merging of spatially adjacent regions is performed. If $0.0 < S_{wght} \leq 1.0$, merging between spatially adjacent regions is favored compared to merging of spatially non-adjacent regions by a factor of $1.0/S_{wght}$. As we have discussed in [139], the optimal parameter S_{wght} can be chosen based on *a priori* knowledge about information classes contained in the image. If some classes have very similar spectral responses, we recommend to choose $S_{wght} = 0.0$, *i.e.*, to perform segmentation only in the spatial domain. Otherwise, we recommend to include the possibility of merging spatially non-adjacent regions, while favoring region growing (for instance, $S_{wght} = 0.1$ can be chosen). If $S_{wght} > 0.0$, labeling of connected components has to be applied after RHSEG in order to obtain a segmentation map where each spatially connected component has a unique label.

RHSEG provides as output a hierarchical sequence of image partitions. In this sequence, a particular object can be represented by several regions at finer levels of details, and can be assimilated with other objects in one region at coarser levels of details. While this hierarchical sequence allows flexibility in choosing the appropriate level of detail for the segmentation map, the manual selection of the appropriate level can be time consuming. Therefore, automatic selection is desirable. The possibility of the automated selection of results for the RHSEG was explored in [113].

2.2.4 Spectral-spatial classification approaches

Once image segmentation is performed, the next step is to incorporate the spatial information defined by regions from a segmentation map in contextual classification. In the following, we propose three major approaches that can be used for accomplishing this step.

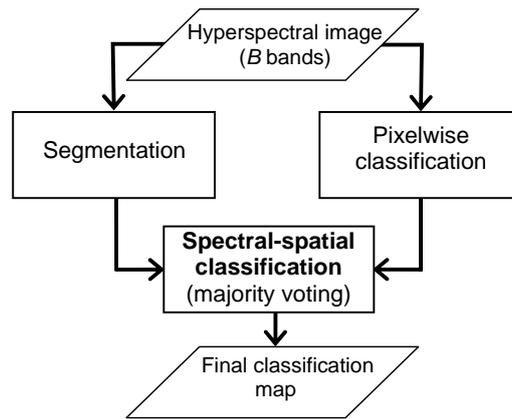


Figure 2.4: Flowchart of the proposed spectral-spatial classification scheme for hyperspectral images using majority voting.

2.2.4.1 Object-based classification

The *object-based classification* approach consists in defining, for every region, a feature or set of features, and then assigning the whole region to one of the classes using this set of features.

Widayati *et al.* [151] and Linden *et al.* [88] applied this approach for spectral-spatial classification of multispectral and hyperspectral remote sensing images, respectively. The most common way to apply the object-based classification is to assign each region from the segmentation map to one of the classes using its vector mean as a feature. The advantages of this approach are simplicity and noise filtering. However, when representing each region by only its vector mean, a significant amount of spectral and textural information about the region may be lost, resulting in imprecisions of classification.

2.2.4.2 Combination of spectral and spatial information within a pixel feature vector

Another type of spectral-spatial classification consists in combining both the spectral and the spatial information within a feature vector of each pixel, and then classifying each pixel using these feature vectors.

Numerous studies explored the use of composite kernels for integration of spectral and spatial information within feature vectors of pixels [21, 22, 44, 45, 78]. The framework of composite kernels, with its application for classification of hyperspectral images, is briefly described in Chapter 1 and more in details in [22]. The authors present four different kernel approaches for the joint consideration of spectral and spatial information into classification: 1) stacked features approach; 2) direct summation kernel; 3) weighted summation kernel; 4) cross-information kernel. The most common method is a *stacked features* approach, in which feature vectors are built from the concatenation of spectral and spatial features.

2.2.4.3 Combination of segmentation and pixelwise classification maps

The third considered context classification approach can be described as follows: A pixelwise classification, based on the spectral information of pixels, and a segmentation are performed independently. Then, the results of a segmentation and classification are combined, using a defined fusion rule.

Li and Xiao [85] used spectral and spatial information for classification of a multispectral (SPOT 5) image. A watershed segmentation and an ML classification were independently performed. Then, pixels of the whole region were assigned to one class if more than 50% of pixels in this region were categorized

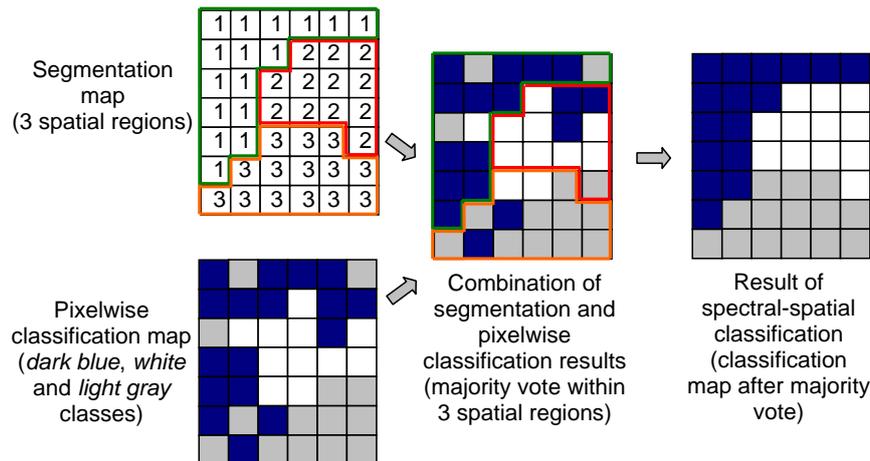


Figure 2.5: Example of spectral-spatial classification using majority voting.

into one class by a pixelwise classifier. The classification results were substantially improved with the spectral-spatial approach compared to the pixelwise classification.

Another approach, called *majority vote*¹, is described in the following. Segmentation and pixelwise classification of an image are independently performed. For every region in the segmentation map, all the pixels are assigned to the most frequent class within this region. Widayati *et al.* [151] used this technique for context classification of multispectral images. Despite its simplicity, this approach has demonstrated the best classification performances when compared to other methods of fusion of spectral and spatial information. We propose to use the *majority voting* approach for context classification of hyperspectral images using segmentation-derived adaptive neighborhoods and SVM (see Papers 2, 3). Figures 2.4 and 2.5 show a flowchart and an example of the combination of spectral and spatial information using the proposed classification method, respectively. An SVM classifier, which efficiently handles hyperspectral data, can therefore be used as a pixelwise classifier in the described classification scheme.

2.2.5 Concluding discussion

A novel spectral-spatial classification approach for hyperspectral images has been presented. We have proposed to perform an image segmentation in order to use every region from the segmentation map as an adaptive neighborhood for all the pixels within this region. Several segmentation techniques based on different principles have been investigated for this purpose and extended to the case of hyperspectral images. Watershed and partitional clustering segmentation techniques are fast and require almost no input parameters, thus enabling automatic image segmentation. The RHSEG method produces a hierarchical sequence of image partitions, thus providing detailed spatial information. However, it is more computationally demanding and the automatic interpretation of the results is more challenging, when compared to the former two techniques. We will further compare these methods in Section 2.4 and will give recommendations on how to choose an appropriate image segmentation technique.

The extracted spatial information must be further combined with spectral information into a classifier. We have presented different approaches for spectral-spatial classification. Finally, we have proposed to use SVM and a majority vote rule for spectral-spatial classification of hyperspectral images. This approach retains all the spectral information for its accurate classification with a well-suited technique, while not increasing data dimensionality. Thus, it has proven to be an accurate, simple and fast technique.

¹In the literature, this approach is often referred to as *plurality vote*.

2.3 Segmentation and Classification Using Automatically Selected Markers

In the previous section, we have explored unsupervised segmentation of hyperspectral images for defining spatial structures. However, unsupervised image segmentation is a challenging task. Segmentation aims at dividing an image into homogeneous regions, but the measure of homogeneity is image-dependent. A too relaxed or a too restricted homogeneity criterion can lead to undersegmented or oversegmented results, respectively. When performing unsupervised segmentation, we preferred oversegmentation to undersegmentation, in order not to miss objects in the classification map (see Papers 2, 3). Our next objective is to reduce oversegmentation, and thus further improve segmentation and classification results. This can be achieved by performing a marker-controlled segmentation [56, 126]. The idea behind this approach is to select for every spatial object one or several pixels belonging to this object (called a *region seed*, or a *marker* of the corresponding region) and to grow regions from the selected seeds, so that every region in the resulting segmentation map is associated with one region seed. In this section, we present new methods for hyperspectral image analysis using marker-controlled region growing. First, we propose three different techniques for automatic marker selection. Then, we develop classification methods using marker-controlled region growing segmentation.

2.3.1 Marker selection

In previous studies, a *marker* (an internal marker) was defined as a connected component belonging to the image and associated with an object of interest [56, 70, 104, 126]. We generalize the definition of a marker as follows:

Definition 2.6 (Marker) *A marker is defined as a set of image pixels (not necessarily connected; it can be composed of several spatially disjoint subsets of connected pixels) which is associated with one object in the image scene.*

The markers of regions can be chosen either manually, which is time consuming, or automatically. The problem of automatic marker selection has previously been discussed in the literature, mostly for grayscale and color images. Markers are often defined by searching flat zones (*i.e.*, connected components of pixels of constant gray level value), zones of homogeneous texture or image *extrema* [126]. Gómez *et al.* [55] applied histogram analysis to obtain a set of representative pixel values, and the markers were generated with all the image pixels with representative gray values. Jalba *et al.* [70] used connected operators filtering on the gradient image, in order to select markers for a grayscale diatom image. Noyel *et al.* [104, 103] performed classification of the hyperspectral image (using different techniques, such as Clara [73] and linear discriminant analysis), and then filtered the classification maps class by class, using morphological operators, in order to select large spatial regions as markers. Furthermore, the authors proposed to use random balls (connected sets of pixels of randomly selected sizes) extracted from these large regions as markers. In the discussed studies [70, 103, 104], the objective was to segment specific structures (blood cells, diatoms, glue occlusions and cancerous growth).

Our objective is to *mark* (select a marker for) each significant spatial object in the image. Here, by significant we mean an object of at least one-pixel size that belongs to one of the classes of interest. As remote sensing images contain small and complex structures, automatic selection of markers is an especially challenging task.

We propose to use probabilistic classification results for choosing the most reliable pixels as markers of spatial regions (Figure 2.6 shows a flowchart of the proposed classification scheme; see Papers 5, 6). Probabilistic classification yields, for each pixel, probability estimates to belong to each class of interest and a unique class label (corresponding to the *maximum* probability estimate, see Section 2.1). From here, we can deduct if, for a given pixel, the assigned class label is reliable (*i.e.*, the probability to actually belong to the assigned class is high) or doubtful. Assuming that classification results are typically more accurate inside spatial regions and more erroneous closer to region borders, we propose to choose the

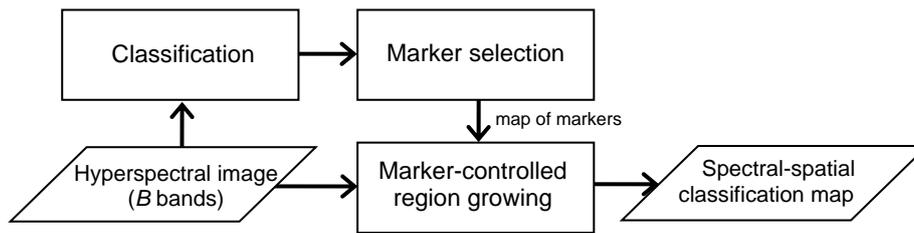


Figure 2.6: Flowchart of the general proposed spectral-spatial classification scheme using marker-controlled region growing.

most reliably classified pixels as region markers. We have developed three different methods of marker selection based on this idea: using ML, SVM classification results and a multiple classifier approach. These three methods are presented in the following sections.

2.3.1.1 Marker selection using ML discriminant values

The proposed marker selection procedure consists of three steps:

1. Feature extraction.
2. Gaussian ML pixelwise classification.
3. Selection of the most reliable classified pixels as markers.

1) Feature extraction. Since ML classification involves estimation of covariance matrices, feature extraction is a required step for reducing the dimensionality of the hyperspectral image. We propose to apply a PCFA method for this purpose [71].

2) Gaussian ML pixelwise classification. As described in Chapter 1, in Gaussian ML classification samples of each class are assumed to be drawn from a multivariate Gaussian distribution (see equation 1.2). Means and covariance matrices $(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ for each class ω_k ($k = 1, \dots, K$) are estimated from the training samples, using ML estimates (see equations 1.3 and 1.4). Next, for each pixel \mathbf{x}_i , K discriminant functions are computed as

$$g_k(\mathbf{x}_i) = \ln p(\omega_k) - \frac{1}{2} \ln |\boldsymbol{\Sigma}_k| - \frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k), \quad k = 1, \dots, K, \quad (2.20)$$

where $p(\omega_k)$ is the prior probability of the class ω_k . When the $p(\omega_k)$ is unknown, a situation of equal prior probabilities is assumed. In this case, the discriminant function can be simplified as follows [117]:

$$g_k(\mathbf{x}_i) = -\ln |\boldsymbol{\Sigma}_k| - (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k), \quad k = 1, \dots, K. \quad (2.21)$$

The pixel \mathbf{x}_i is assigned to the class ω_k , if $g_k(\mathbf{x}_i) > g_j(\mathbf{x}_i)$ for all $j \neq k$. The corresponding value of a discriminant function $g_k(\mathbf{x}_i)$ (*winning discriminant value*) is stored in a so called *discriminant map* $\mathcal{D} = \{d_j \in \mathbb{R}, j = 1, \dots, n\}$, where n is a number of pixels: $d_i = g_k(\mathbf{x}_i)$.

3) Selection of the most reliable classified pixels as markers. We propose to use a discriminant map obtained as a result of the pixelwise classification in order to keep the most reliably classified pixels as markers. For a pixel \mathbf{x} assigned to the class ω_k , the higher value of $d = g_k(\mathbf{x})$ corresponds to the higher probability of a correct classification. Thus, by comparing winning discriminant values of two pixels assigned to the same class, we can conclude which pixel is more reliably classified. However, winning discriminant values for pixels assigned to different classes are not comparable (this can be deduced from the Bayes theorem [40] and equation 2.21). Consequently, we propose to define for each class ω_k a threshold of classification reliability $\tau(\omega_k)$, and then keep all the reliably classified pixels as markers. The proposed marker selection procedure consists of the following steps:

- When performing pixelwise classification, find the *maximum* and *minimum* values of the discriminant functions for each class ω_k :

$$g_k^{MAX} = \max_{\mathbf{x} \in \mathbf{X}} g_k(\mathbf{x}), \quad (2.22)$$

$$g_k^{MIN} = \min_{\mathbf{x} \in \mathbf{X}} g_k(\mathbf{x}). \quad (2.23)$$

- For each class ω_k , compute a threshold of classification reliability:

$$\tau(\omega_k) = g_k^{MAX} - t(g_k^{MAX} - g_k^{MIN}), \quad 0 < t < 1, \quad (2.24)$$

where t is the parameter that controls the threshold level. The higher the value of this parameter is, the more pixels are retained as markers.

- For every pixel \mathbf{x}_i , apply the following procedure: if it is assigned to the class ω_k and $d_i \geq \tau(\omega_k)$, keep this pixel as a marker.

In the resulting map of markers, each marker pixel is associated with the class defined by the pixelwise classifier. The number of markers can be further reduced, by merging either spatially connected markers with the same class label, or markers issued from the same connected component (*i.e.*, spatially connected set of pixels assigned to the same class) in the pixelwise classification map. In the latter case, a marker is not necessarily a connected set of pixels: It can spatially be split into several subsets.

The main advantage of the described marker selection technique is its algorithmic and computational simplicity. However, this method has the following disadvantages: 1) The ML classifier is not well adapted for classifying high-dimensional data, when a limited number of training samples is available. 2) The proposed thresholding method does not take the spatial information into account when selecting marker pixels. For a given pixel \mathbf{x} assigned to the class ω_k , the probability of a correct classification is higher if its neighbors are assigned to the same class. Therefore, it is of interest to consider spatial dependencies between pixels during the marker selection procedure.

2.3.1.2 Marker selection using probabilistic SVM

To cope with the drawbacks of the previously described approach, we have proposed another method for marker selection (see Paper 5). The new technique differs from the previous one mainly in two ways: 1) SVM probabilistic classification is used for computing the classification map and probabilities of a correct classification. 2) When performing marker selection by selecting the most reliable classified pixels, each connected component in the classification map is analyzed, instead of the pixel by pixel analysis. The proposed marker selection method consists of two steps:

1. Probabilistic pixelwise classification.
2. Selection of the most reliable classified pixels as markers.

1) Probabilistic pixelwise classification. At this step, it is desirable to choose a classifier that can efficiently handle high-dimensional data. We propose to use an SVM classifier for this purpose [110, 148]. The probabilistic SVM classification is described in Section 2.1.1. The outputs of this step are the following:

1. classification map, where each pixel has a unique class label;
2. **probability map**, containing probability estimates for each pixel to belong to the assigned class.

Probabilistic SVM classification yields, for each pixel \mathbf{x} , probability estimates to belong to each class of interest $\mathbf{p} = \{p_k = p(y = k|\mathbf{x}), k = 1, \dots, K\}$. A probability map is constructed by assigning to each pixel the maximum probability estimate $\max(p_k), k = 1, \dots, K$.

2) Selection of the most reliable classified pixels as markers. We propose to use probabilistic classification results in order to choose the most reliably classified pixels as markers. The proposed marker selection procedure can be described as follows (see the flowchart in Figure 2.7(a)):

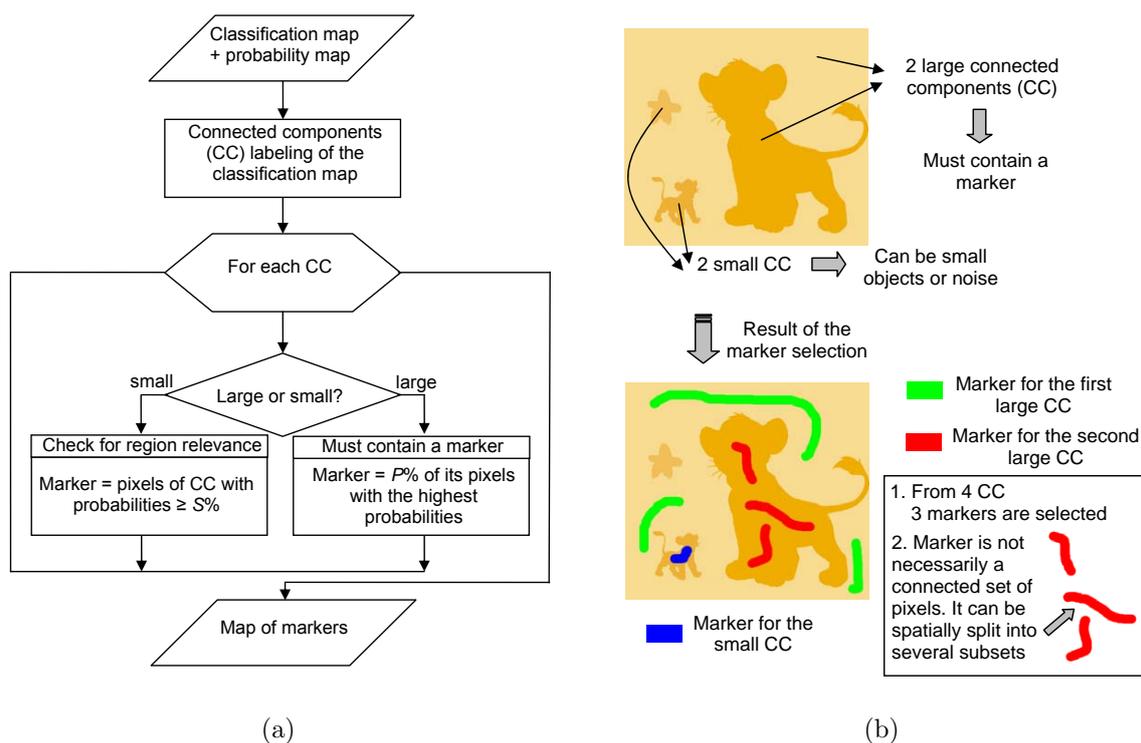


Figure 2.7: (a) Flowchart of the proposed marker selection procedure. (b) Illustrative example of the marker selection.

1. Perform a *connected components labeling* of the pixelwise classification map (we propose to use an eight-neighborhood connectivity). For this purpose, a classical connected-component algorithm using the union-find data structure can be used [123].
2. Analyze each *connected region* as follows:
 - If a region is *large* enough, it *should contain a marker*, which is determined as $P\%$ of the pixels within the connected component with the highest probability estimates.
 - If a region is *small*, it should lead to a marker *only if it is very reliable*; a potential marker is formed by pixels with probability estimates higher than a defined threshold.

The proposed procedure is deduced from the following analysis: Based on the results of our previous studies (see Papers 2, 3) [136, 137, 140], it is common that almost no undersegmentation is present in a pixelwise classification map. Therefore, each connected spatial region from the classification map is analyzed if it corresponds most probably to the spatial object or if it is rather a classification noise [an illustrative example is given in Figure 2.7(b)]. If the size of the component is large enough to consider it as a relevant region, the most reliably classified pixels within this region are selected as its marker. If a component contains only a few pixels, it is investigated if these pixels were classified to a particular class with a high probability. If this is the case, the considered connected component represents a small spatial structure. Thus, a marker associated with this region must be defined. Otherwise, the component is the consequence of classification noise, and we tend to eliminate it. Therefore, no marker within this component is selected.

For the proposed marker selection procedure, the following parameters must be chosen:

- A parameter M defining if a region is considered as being large or small. We propose to use a number of pixels in the region as a criterion of the region size. The threshold of the number of pixels defining if the region is large depends on the resolution of the image and typical sizes of the objects of interest.

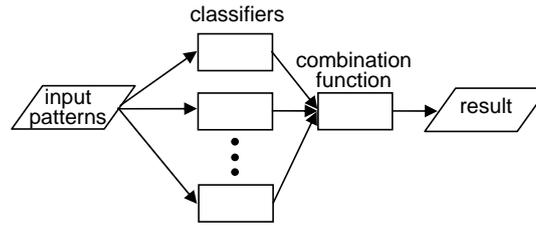


Figure 2.8: Schematic diagram of a multiple classifier system.

- A parameter P , defining the percentage of pixels within the large region to be used as markers, depends on the previous parameter. Since a marker for the large region must be composed at least of one pixel, the following condition must be fulfilled: $P \geq 100\%/M$.
- The last parameter S , which is a threshold of probability estimates defining potential markers for small regions, depends on the probability of the presence of small structures in the image (which depends on the image resolution and the classes of interest), and the importance of the potential small structures (*i.e.*, what is the cost of losing the small structures in the classification map). This classification reliability threshold can be set either equal for all the classes, or class-specific thresholds can be chosen. The latter approach can yield more accurate results, if thresholds are accurately chosen. However, the former approach may be preferable since it is computationally simpler.

In Paper 5 we have given recommendations for the choice of parameters, and have investigated experimentally the dependence of the classification accuracies from the chosen parameters. As a conclusion, each connected set of pixels with the same class in the classification map provides either one or zero marker. Each marker can be composed of several spatially disjoint subsets of adjacent pixels, and each marker has a class label.

2.3.1.3 Multiple classifier approach

Although the previously described marker selection approach has shown excellent results, the drawback of this method is that the choice of markers strongly depends on the performances of the selected pixelwise classifier (e.g., the SVM classifier). Our next objective is to mitigate the dependence of the marker selection procedure from the choice of a pixelwise classifier. This can be achieved by using not a single classification algorithm for marker selection, but rather an ensemble of classifiers, *i.e.*, *multiple classifiers*. For this purpose, several individual classifiers must be chosen and combined within one system in such a way that the complementary benefits of each classifier are used, while their weaknesses are avoided [17, 75, 153].

We propose a new marker selection method based on a *Multiple Classifier (MC)* system (see Paper 6). A schematic representation of an MC system is given in Figure 2.8. The proposed marker selection approach consists of two steps:

- Multiple classification: Several classifiers are used independently to classify an image.
- Marker selection: A marker map is constructed by selecting the pixels assigned by all the classifiers to the same class.

1) Multiple classification. At this step, several individual classifiers are applied to an image. An important issue for an efficient MC system is that the individual classifiers should be independent. More precisely, the classifiers should not agree with each other when they misclassify a pixel [75]. The complementary properties of the different classifiers selected for the MC system should ensure to a certain extent this requirement. For instance, standard pixelwise classification algorithms can be used for

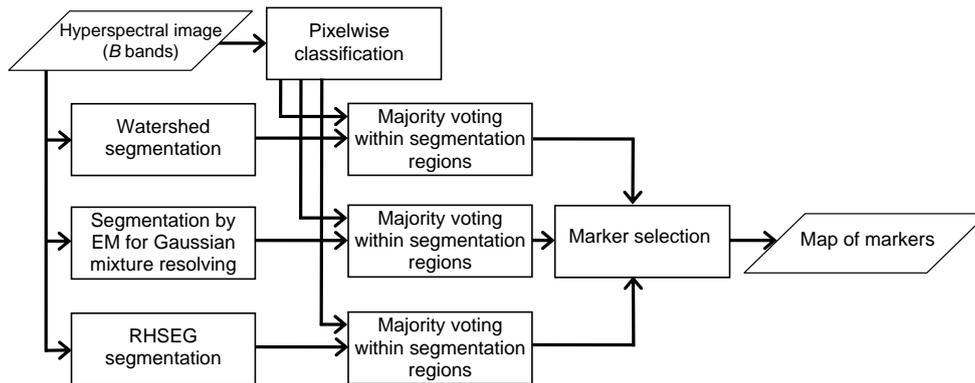


Figure 2.9: Flowchart of the proposed *MSSC* marker selection scheme.

this purpose, such as SVM, ML and k-NN methods (parametric and non-parametric techniques, based on different principles). We use these individual techniques in what we call the *MC* marker selection scheme.

Furthermore, we propose to use spectral-spatial classifiers as individual classifiers for the MC system (*Multiple Spectral-Spatial Classifier (MSSC)* marker selection scheme), each of them combining the results of a pixelwise classification and one of the unsupervised segmentation techniques. Figure 2.9 shows a flow-chart of the proposed *MSSC* marker selection scheme, which consists of the following steps:

1. *Unsupervised image segmentation*: Segmentation methods based on different principles must be chosen. We have investigated the use of three techniques: Watershed segmentation, segmentation by EM and segmentation using the RHSEG method. These methods and their extension to the case of hyperspectral images are described in Section 2.2 and in Papers 2, 3 and 6.
2. *Pixelwise classification*: We propose to use an SVM method for classifying a hyperspectral image (see Chapter 1). This step results in a classification map, where each pixel has a unique class label.
3. *Majority voting within segmentation regions*: Each of the obtained unsupervised segmentation maps is combined with the pixelwise classification map using the majority voting principle: For every region in the segmentation map, all the pixels are assigned to the most frequent class within this region (see Section 2.2.4). Thus, q segmentation maps combined with the pixelwise classification map result in q spectral-spatial classification maps (since we propose to use three different segmentation techniques, in this particular case $q = 3$).

Different segmentation methods based on dissimilar principles lead to different classification results. By using spectral-spatial classifiers in this step, spatial context in the image is taken into account, yielding more accurate classification maps when compared to those obtained by performing pixelwise classification.

2) Marker selection. Another important issue for designing an MC system is the rule for combining the individual classifiers (*i.e.*, combination function). The individual classifier outputs, such as class labels and possibly posterior probabilities, are typically combined by voting rules, belief functions, statistical techniques, the Dempster-Shafer evidence theory, and other schemes [153]. For a given pixel, if all the classifiers agree on the same class k , the evident combination rule consists in assigning this pixel to the class k in the final classification map. On the other side, when individual classifiers disagree in assigning the given pixel, the procedure of final decision making is not that straightforward, and different combination functions may yield different results. A typical result of the MC system is a final classification map, where each pixel has a unique class label. This type of MS systems has previously been used for remote sensing image classification [17, 49, 146].

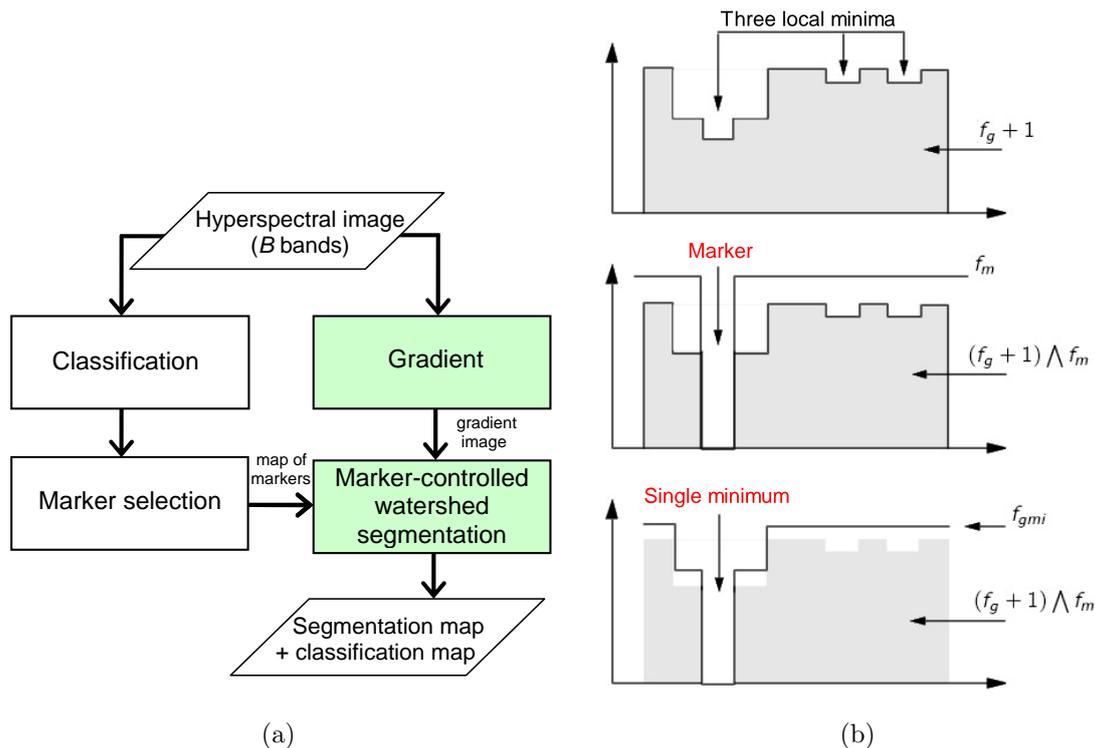


Figure 2.10: (a) Flowchart of the proposed spectral-spatial classification method using marker-controlled watershed. (b) Illustrative example of the *minima* imposition.

We propose to address the combination rule issue in the following way. A map of markers is computed, using classification maps from the previous steps and exclusionary rule: For every pixel, if all the classifiers agree, the pixel is kept as a marker, with the corresponding class label. The resulting map of m markers contains the most reliably classified pixels. The rest of the pixels are further classified by performing a marker-controlled region growing, as described in the following section.

2.3.2 Classification using marker-controlled region growing

Once marker selection is performed, the obtained map of markers is further used for marker-controlled region growing. We have proposed two different methods for this purpose: marker-controlled watershed and construction of a Minimum Spanning Forest (MSF). We present these two methods in the following sections.

2.3.2.1 Marker-controlled watershed

The first proposed technique is based on the standard marker-controlled watershed segmentation algorithm, using minima imposition followed by flooding simulations [126] (see Paper 4). This approach determines markers for each region of interest and transforms the gradient image in such a way that the region markers are the only local *minima* of the resulting image. Figure 2.10(a) depicts a flowchart of the proposed spectral-spatial classification method using marker-controlled watershed. In the following, we describe the marker-controlled region growing part of this scheme (corresponding to the green boxes in the figure), which consists of two steps:

1) Gradient. A one-band gradient $f_g = \nabla(\mathbf{X})$ of a hyperspectral image is computed, which is needed as an input for a watershed segmentation. Approaches for computing a one-band gradient from

the multi-band image are discussed in Section 2.2.1.

2) **Marker-controlled watershed segmentation.** In order to incorporate the use of markers into the watershed algorithm, we use the *minima* imposition technique described in [126]. We propose the following marker-controlled segmentation scheme:

1. First, the marker image is created:

$$f_m(\mathbf{x}) = \begin{cases} 0, & \text{if pixel } \mathbf{x} \text{ belongs to a marker,} \\ t_{max}, & \text{otherwise.} \end{cases} \quad (2.25)$$

Before creating the marker image, it must be checked that each marker is spatially disconnected from any other marker (it is especially important to verify that markers with different class labels do not “touch” each other).

2. Then, the *minima* imposition technique is applied to the gradient image f_g [see an illustrative example of the *minima* imposition in Figure 2.10(b)]. The resulting image f_{gmi} is computed as a morphological reconstruction by erosion of $(f_g + 1) \wedge f_m$ (point-wise *minimum* between the gradient image and the marker image) from the marker image f_m :

$$f_{gmi} = R_{(f_g+1)}^\varepsilon \wedge_{f_m} (f_m) \quad (2.26)$$

3. The algorithm of Vincent and Soille [149] for watershed is applied on the filtered gradient image f_{gmi} . In the resulting segmentation map, each pixel contains either a label of the region it belongs to, or a watershed pixel label. Here, region labels do not correspond to marker labels. The Vincent and Soille algorithm creates a new region label for each local *minimum*, and then grows regions from *minima*. In our case, a marker can contain non-adjacent groups of pixels. After the *minima* imposition, this will lead to several local *minima* in the filtered gradient image. Therefore, each marker leads to one or several regions.
4. In order to obtain a segmentation map without border pixels, each watershed pixel is assigned to the neighboring region with the “closest” median, *i.e.*, the distance between the vector median of this region and the watershed pixel vector is minimal (see Section 2.2.1).
5. Regions belonging to the same marker are merged together and are associated with this marker.
6. Finally, pixels of each region are assigned to the class of its marker. This results in a spectral-spatial classification map.

2.3.2.2 Construction of a Minimum Spanning Forest

The second proposed marker-controlled segmentation technique consists in the grouping of all the image pixels into an **MSF** [130], where each tree is rooted on a classification-derived marker² (Figure 2.11 shows a flow-chart of the classification scheme using the proposed approach; see Paper 5). The decision to connect the pixel, which is not yet in the forest, to one of the trees in the forest is based on its similarity to one of the adjacent pixels already belonging to the forest. By assigning a class of the marker to all the pixels within the region grown from the considered marker, a spectral-spatial classification map is obtained. Furthermore, the classification map is refined, using results of a pixelwise classification and a majority voting within the spatially connected regions [137].

The construction of an MSF belongs to graph-based approaches for image segmentation [100, 124, 130, 135]. They introduce the Gestalt principles of perceptual grouping to the field of computer vision. The image is associated with a graph, the vertices of which correspond to the image entities (pixels or

²In the recent works, extensions and generalizations of a watershed segmentation method have been proposed [5, 32]. In particular, it was shown that a watershed segmentation can also be built from an MSF [32, 130]. We would like to stress that in our work, we consider the watershed segmentation using the classical paradigm of the morphological image segmentation: a gradient of the image is computed; then, a minima imposition technique is applied, followed by the watershed algorithm. The approach, proposed in this section, based on the construction of an MSF using classification-derived markers and an arbitrary dissimilarity measure is a general, simple and efficient region growing segmentation technique.

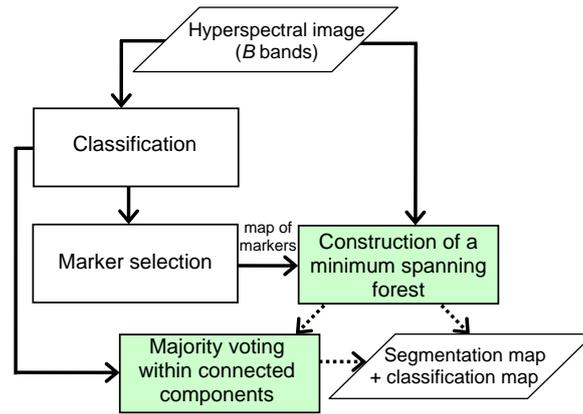


Figure 2.11: Flowchart of the proposed spectral-spatial classification approach using an MSF grown from automatically selected markers.

regions) and the edges correspond to relations between these entities. A weight associated with each edge indicates the (dis)-similarity between two entities (pixels or regions). In [97], Meyer proposed the use of an MSF for image segmentation. Several recent publications further investigate this topic [5, 32, 130]. However, the authors of these works do not explore the problem of the automatic marker selection. Their segmentation is based on markers provided by the user.

In the following, the two steps of the proposed classification procedure based on marker-controlled region growing are described:

1. Construction of an MSF.
2. Majority voting within connected components.

1) Construction of an MSF. In this step, all the image pixels are grouped into an MSF [130], where each tree is rooted on a classification-derived marker. For this purpose, each pixel is considered as a vertex $v \in V$ of an undirected graph $G = (V, E, W)$, where V and E are the sets of vertices and edges, respectively, and W is a mapping of the set of the edges E into \mathbb{R}^+ . Each edge $e_{i,j} \in E$ of this graph connects a couple of vertices i and j corresponding to the neighboring pixels (in the following, we simply call vertices as pixels). If an 8-neighborhood is used, every pixel is connected by an edge with each of its neighbors (in total 8 edges). Furthermore, a weight $w_{i,j}$ is assigned to each edge $e_{i,j}$, which indicates the degree of dissimilarity between two pixels connected by this edge. Different dissimilarity measures can be used for computing weights of edges, such as vector norms, SAM (see equation 1.5), Spectral Information Divergence (SID) [93].

Definition 2.7 (L1 vector norm) The L1 vector norm between two pixel vectors $\mathbf{x}_i = (x_{i1}, \dots, x_{iB})^T$ and $\mathbf{x}_j = (x_{j1}, \dots, x_{jB})^T$ is given as

$$L1(\mathbf{x}_i, \mathbf{x}_j) = \sum_{b=1}^B |x_{ib} - x_{jb}|. \quad (2.27)$$

Definition 2.8 (SID measure) The SID measure [26] computes the discrepancy of probabilistic behaviors between the spectral signatures of two pixels. It is defined as

$$SID(\mathbf{x}_i, \mathbf{x}_j) = \sum_{b=1}^B \left\{ q_b(\mathbf{x}_i) \log \left[\frac{q_b(\mathbf{x}_i)}{q_b(\mathbf{x}_j)} \right] + q_b(\mathbf{x}_j) \log \left[\frac{q_b(\mathbf{x}_j)}{q_b(\mathbf{x}_i)} \right] \right\}, \quad (2.28)$$

where

$$q_b(\mathbf{x}_i) = \frac{x_{ib}}{\sum_{l=1}^B x_{il}}. \quad (2.29)$$

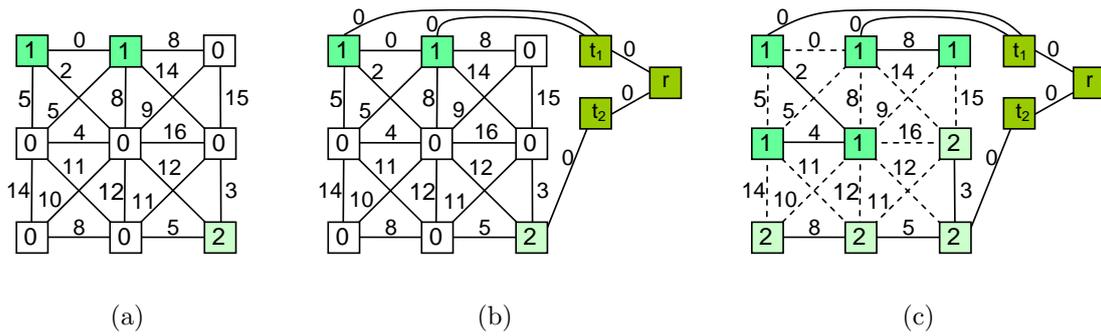


Figure 2.12: Example of construction of an MSF rooted on markers. (a) The original image graph G , where colored vertices represent markers 1 and 2; non-marker pixels are denoted by “0”. (b) Addition of extra vertices t_1, t_2, r to the graph. (c) Minimum spanning tree of the graph presented in (b); after removing the vertex r , an MSF is obtained, where each tree grown from the vertex t_i forms a region in the segmentation map.

Furthermore, more complex dissimilarity measures for image segmentation have been proposed in [31, 90].

Definition 2.9 (Spanning tree) Given a connected graph $G = (V, E)$, a spanning tree $T = (V, E_T)$ of G is a connected graph without cycles such that $E_T \subset E$.

Definition 2.10 (Spanning forest) Given a connected graph $G = (V, E)$, a spanning forest $F = (V, E_F)$ of G is a non-connected graph without cycles such that $E_F \subset E$.

Definition 2.11 (Minimum spanning tree) Given a graph $G = (V, E, W)$, the minimum spanning tree is defined as a spanning tree $T^* = (V, E_{T^*})$ of G such that the sum of the edges weights of T^* is minimal:

$$T^* \in \arg \min_{T \in ST} \left\{ \sum_{e_{i,j} \in E_T} w_{i,j} \right\}, \quad (2.30)$$

where ST is a set of all spanning trees of G .

Definition 2.12 (Rooted Minimum Spanning Forest) Given a graph $G = (V, E, W)$, the MSF rooted on a set of m distinct vertices $\{t_1, \dots, t_m\}$ consists in finding a spanning forest $F^* = (V, E_{F^*})$ of G , such that each distinct tree of F^* is grown from one root t_i , and the sum of the edges weights of F^* is minimal [130]:

$$F^* \in \arg \min_{F \in SF} \left\{ \sum_{e_{i,j} \in E_F} w_{i,j} \right\}, \quad (2.31)$$

where SF is a set of all spanning forests of G rooted on $\{t_1, \dots, t_m\}$.

In order to obtain the MSF rooted on markers, m additional vertices $t_i, i = 1, \dots, m$ are introduced. Each extra vertex t_i is connected by the edge with a null weight to the pixels representing a marker i . Furthermore, an additional root vertex r is added and is connected by the null-weight edges to the vertices t_i . The minimum spanning tree of the constructed graph induces an MSF in G , where each tree is grown on a vertex t_i ; the MSF is obtained after removing the vertex r . An example of construction of the MSF rooted on markers is shown in Figure 2.12. Prim’s algorithm can be used for building the MSF (see Algorithm 1 in Paper 5) [115]. The efficient implementation of the algorithm using a binary min-heap (for the implementation of a min-priority queue) is possible [30], the resulting time complexity of the algorithm is $O(|E| \log |V|)$.

Each tree in the MSF forms a region in the segmentation map (by mapping the resulting graph onto an image). Finally, a spectral-spatial classification map is obtained by assigning the class of each marker to all the pixels grown from this marker.

Thus, the proposed procedure of the construction of an MSF from region markers is a region growing method, which consists of the following steps: First, seed regions are chosen to belong to the segmentation and classification maps. Then, at each iteration a new pixel i is added to the segmentation and classification maps, so that the dissimilarity criterion between this pixel and one of the pixels j already belonging to the segmentation/classification map is minimal. When including the new pixel to the classification map, a class of the pixel j is assigned to the pixel i .

2) Majority voting within connected components (optional step). Although the most reliable classified pixels are selected as markers, it may happen that a marker is classified to the wrong class. In this case, all the pixels within the region grown from this marker risk to be wrongly classified. In order to make the proposed classification scheme more robust, we propose to post-process the classification map, by applying a simple *majority voting* technique which has shown good performances for spectral-spatial classification [137, 141, 151]. For this purpose, connected component labeling is applied on the obtained spectral-spatial classification map (using a four-neighborhood connectivity). Furthermore, for every connected component (region), all the pixels are assigned to the most frequent class when analyzing a pixelwise classification map within this region.

Note that we propose to use an eight-neighborhood connectivity for the construction of an MSF and a four-neighborhood connectivity for the majority voting. The use of the eight-neighborhood connectivity in the first case enables one to obtain more accurate segmentation map, without rough borders. Since an MSF is built from the set of markers, the number of regions does not depend on the chosen connectivity. When performing the last majority voting step, the use of the four-neighborhood connectivity results in the larger or the same number of connected components as the use of the eight-neighborhood connectivity. Therefore, the possible undersegmentation can be corrected in this step. One region from a segmentation map can be split into two connected regions when using the four-neighborhood connectivity. Furthermore, these two regions can be assigned to two different classes by the majority voting procedure.

2.3.3 Concluding discussion

Novel methods for spectral-spatial classification of hyperspectral images using region growing segmentation based on automatically derived markers have been presented. The general proposed approach consists in analyzing probabilistic classification results for selecting the most reliably classified pixels as markers of spatial regions. The selected markers are further used for region growing segmentation which results in both segmentation and spectral-spatial classification maps.

Three different marker selection methods have been proposed, using ML, SVM classification results and a multiple classifier approach. When comparing these approaches, several conclusions can be made: 1) It is interesting to use an SVM classifier, which is well suited for handling hyperspectral data, in the marker selection procedure. 2) Spatial information should be taken into account when selecting markers. 3) In order to mitigate the dependence of the marker selection procedure from the choice of a classifier, it is of interest to use multiple classifier approaches.

Next, two different methods for marker-controlled region growing have been proposed: marker-controlled watershed and construction of an MSF. While marker-controlled watershed requires computation and filtering of the one-band gradient and does not handle the case of spatially adjacent markers, the method based on the construction of an MSF using an arbitrary dissimilarity measure is a general, simple, flexible and efficient region growing segmentation technique. Further conclusions about performances of the proposed techniques will be done in the following section, where experimental evaluation of the proposed methods is presented.

2.4 Experimental Evaluation

This section presents results of experimental evaluation and comparison of the proposed approaches for spectral-spatial classification of hyperspectral data. Four hyperspectral airborne images are used to demonstrate experimental results, with different contexts (agricultural, volcano, and urban areas) and acquired by different sensors (AVIRIS and ROSIS airborne imaging spectrometers). Here, classification performances are assessed using:

1. Classification accuracy measures: overall accuracy, class-specific accuracy, average accuracy, and kappa coefficient.
2. Visual comparison of classification maps.

Classification accuracy is estimated by evaluating the exactitude of a given classification map as compared to the reference map (a set of manually labeled pixels). Based on this evaluation, a **confusion matrix** is typically constructed.

Definition 2.13 (Confusion matrix) *In the field of supervised learning, a confusion matrix is a table, where each column represents the instances in a predicted class, while each row represents the instances in an actual class. One benefit of a confusion matrix is that it is easy to see where the system is confusing (i.e., commonly mis-labelling one class as another).*

An example of the confusion matrix is given in Table 2.1 for a three-class problem. C_i represents the class i and C_{ij} is the number of pixels classified to the class j and referenced as the class i . The measures of classification accuracy can be further computed from the confusion matrix, as described hereafter.

Definition 2.14 (Overall Accuracy) *The Overall Accuracy (OA) is the percentage of correctly classified pixels (K is the number of classes):*

$$OA = \frac{\sum_i^K C_{ii}}{\sum_{ij} C_{ij}} \times 100\%. \quad (2.32)$$

Definition 2.15 (Class Accuracy) *The Class-Specific Accuracy (or producer's accuracy) (CA) is the percentage of correctly classified pixels for a given class.*

$$CA_i = \frac{C_{ii}}{\sum_j^K C_{ij}} \times 100\%. \quad (2.33)$$

Definition 2.16 (Average Accuracy) *The Average accuracy (AA) is the mean of class-specific accuracies for all the classes.*

$$AA = \frac{\sum_i^K CA_i}{K} \times 100\%. \quad (2.34)$$

Definition 2.17 (Kappa Coefficient) *The Kappa Coefficient (κ) is the percentage of agreement, i.e., correctly classified pixels, corrected by the number of agreements that would be expected purely by chance. It is generally thought to be a more robust measure than simple percent agreement calculation since κ takes into account the agreement occurring by chance.*

$$\begin{aligned} \kappa &= \frac{P_o - P_e}{1 - P_e} \times 100\%, \\ P_o &= OA/100\%, \\ P_e &= \frac{1}{N^2} \sum_i^K C_{i\cdot} C_{\cdot i}, \\ C_{i\cdot} &= \sum_j^K C_{ij}, \\ C_{\cdot i} &= \sum_j^K C_{ji}, \end{aligned} \quad (2.35)$$

Table 2.1: *Confusion Matrix, N is the Number of Referenced Pixels and K is the Number of Classes.*

Percentage	Classification data				
Reference data	C_1	C_2	C_3	Row total	Class-specific accuracy
C_1	C_{11}	C_{12}	C_{13}	$\sum_i^K C_{1i}$	$\frac{C_{11}}{\sum_i^K C_{1i}}$
C_2	C_{21}	C_{22}	C_{23}	$\sum_i^K C_{2i}$	$\frac{C_{22}}{\sum_i^K C_{2i}}$
C_3	C_{31}	C_{32}	C_{33}	$\sum_i^K C_{3i}$	$\frac{C_{33}}{\sum_i^K C_{3i}}$
Column total	$\sum_i^K C_{i1}$	$\sum_i^K C_{i2}$	$\sum_i^K C_{i3}$	N	
User's accuracy	$\frac{C_{11}}{\sum_i^K C_{i1}}$	$\frac{C_{22}}{\sum_i^K C_{i2}}$	$\frac{C_{33}}{\sum_i^K C_{i3}}$		

where N is the number of referenced pixels.

The high values of the OA, CA, AA, and κ measures (close to 100%) mean that the classification results are accurate. More detailed experimental evaluation of the approaches developed in the frame of this thesis can be found in Part II. In particular, McNemar's test has been performed for evaluating the statistical significance of the differences in classification accuracies obtained for different classification maps (Papers 1, 5) [47]. Furthermore, the accuracy of segmentation and marker selection results has been quantitatively and qualitatively assessed (Papers 3-5, [139]). In the following, four hyperspectral data sets used for validation of the proposed techniques are briefly described. Then, experimental results are presented and discussed.

2.4.1 Data sets

Experimental results are demonstrated on four hyperspectral airborne images recorded by the AVIRIS and the ROSIS sensors, with different contexts (agricultural, volcano, and urban areas), different spatial resolutions (1.3 m and 20 m) and different numbers of spectral channels (from 102 to 200 bands). These four data sets are detailed in the following:

1. The *Indian Pines* image is of a vegetation area that was recorded by the AVIRIS sensor over the Indian Pines test site in Northwestern Indiana. The image has spatial dimensions of 145 by 145 pixels, and a spatial resolution of 20 m per pixel. Twenty water absorption bands have been removed [134], and a 200-band image was used for the experiments. Sixteen classes of interest are considered, which are detailed in Table 2.2, with a number of samples for each class in reference data. A three-band false color image and the reference data are presented in Figure 2.13. We have randomly chosen 50 samples for each class from the reference data as training samples, except for classes "alfalfa", "grass/pasture-mowed" and "oats". These classes contain a small number of samples in the reference data. Therefore, only 15 samples for each of these classes were randomly chosen to be used as training samples. The remaining samples composed the test set.
2. The *Hekla* image was acquired by the AVIRIS sensor over the region surrounding the central-volcano Hekla in Iceland [10]. The AVIRIS sensor operates in the wavelength range from 0.4 μm to 2.4 μm , and utilizes four spectrometers collecting 224 data channels. During the data collection, spectrometer four was not working properly. The 64 data channels recorded by this spectrometer were deleted from the data, along with the first channels for the other three spectrometers (those channels were blank). Therefore, the 157 remaining data channels were used for experiments. The considered image has spatial dimensions of 560 by 600 pixels. Twelve land cover classes of

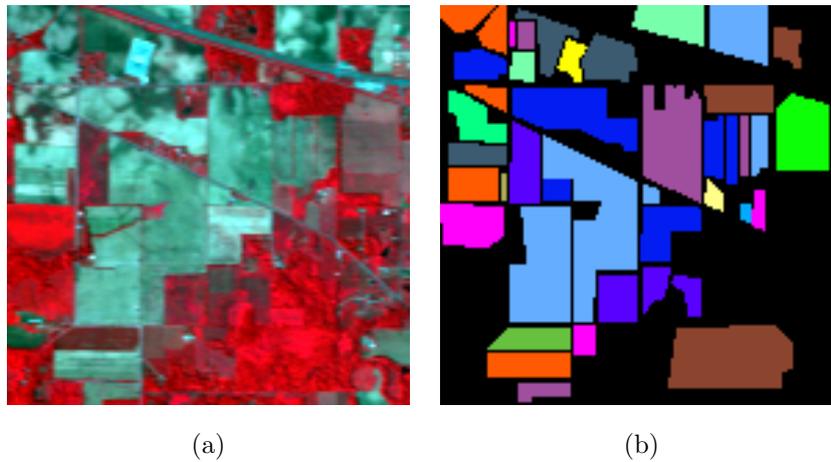


Figure 2.13: *Indian Pines* image. (a) Three-band color composite (837, 636, and 537 nm). (b) Reference data: Corn-no till, Corn-min till, Corn, Soybeans-no till, Soybeans-min till, Soybeans-clean till, Alfalfa, Grass/pasture, Grass/trees, Grass/pasture-mowed, Hay-windrowed, Oats, Wheat, Woods, Bldg-grass-tree-drives, and Stone-steel towers.

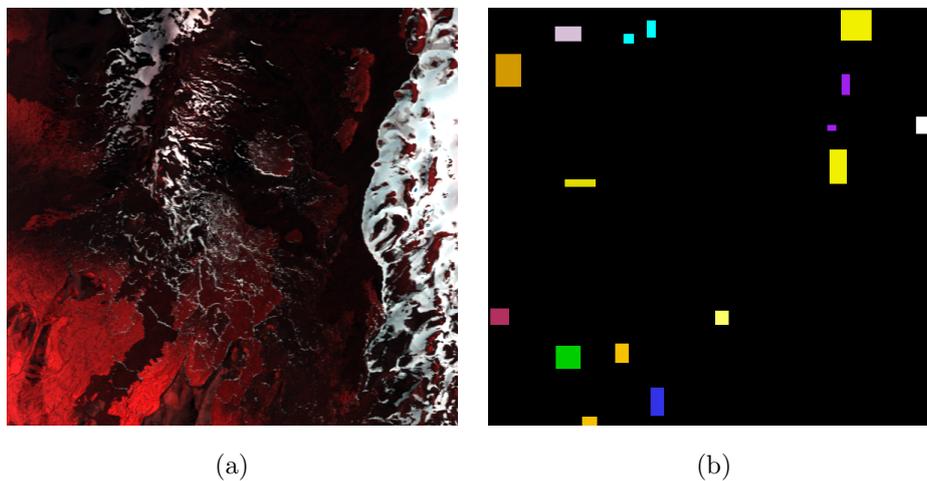


Figure 2.14: *Hekla* image. (a) Three-band color composite (1125, 636, and 567 nm). (b) Reference data: Andesite lava 1970, Andesite lava 1980 I, Andesite lava 1980 II, Andesite lava 1991 I, Andesite lava 1991 II, Andesite lava with moss cover, Hyaloclastite formation, Lava with tephra and scoria, Rhyolite, Scoria, Firn and glacier ice, and Snow (white).

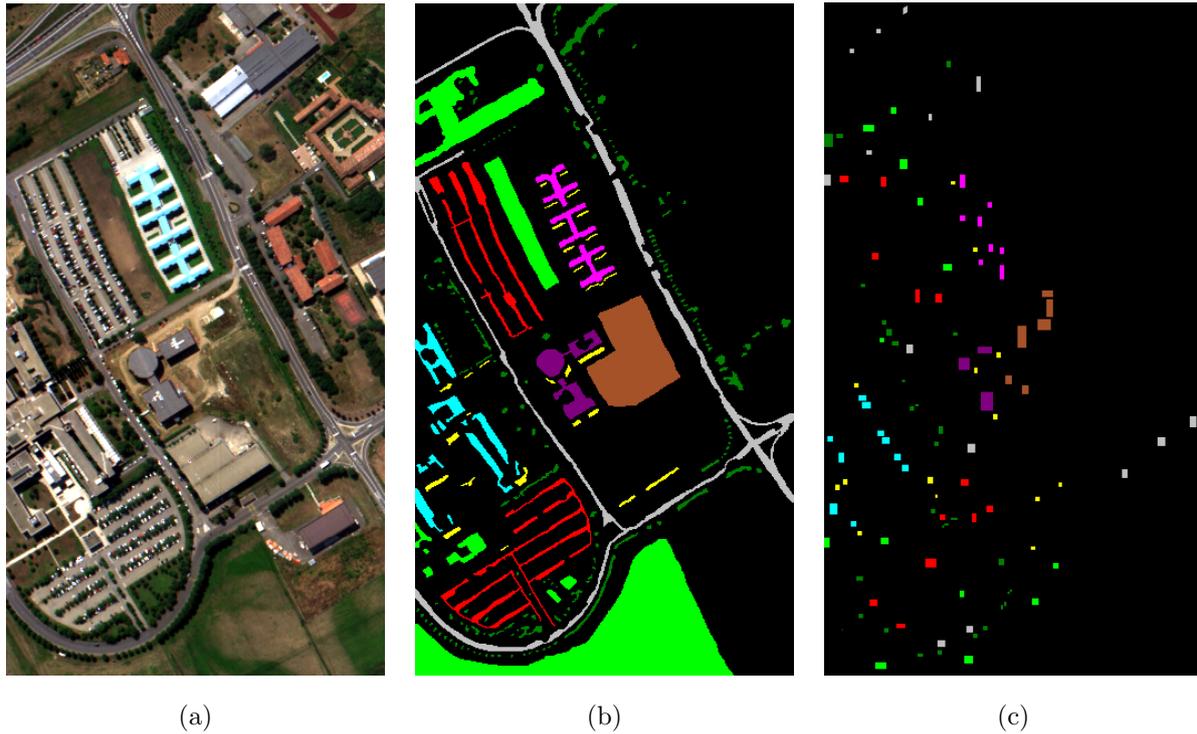


Figure 2.15: *University of Pavia* image. (a) Three-band color composite (650, 558, and 478 nm). (b) Reference data: Asphalt, Meadows, Gravel, Trees, Metal sheets, Bare soil, Bitumen, Bricks, and Shadows. (c) Training data.

interest are considered, which are detailed in Table 2.4, with a number of labeled samples for each class. Figure 2.14 depicts a three-band false color image and the reference data. Fifty samples for each class were randomly chosen from the reference data as training samples, and the rest of the samples were used as the test set.

3. The *University of Pavia* image was recorded by the ROSIS optical sensor over the urban area of the University of Pavia, Italy. The flight was operated by the Deutschen Zentrum für Luft- und Raumfahrt (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Union. The image is 610 by 340 pixels, with a spatial resolution of 1.3 m/pixel. The number of data channels in the acquired image is 115 (with a spectral range from 0.43 to 0.86 μm). The 12 most noisy channels have been removed, and the remaining 103 bands were used for the experiments. The reference data contain nine classes of interest, which are detailed in Table 2.6, with the number of training and test samples for each class. Figure 2.15 shows a three-band false color image, and the reference and training data. A set of training samples provided with the image was used for the experiments.
4. The *Center of Pavia* image is of an urban area that was recorded by the ROSIS sensor during the same flight campaign as the *University of Pavia* data set. The image used for the experiments is 900 by 300 pixels, with 102 spectral channels (the 13 most noisy channels have been removed). The reference data contain nine classes of interest, detailed in Table 2.8, with a number of labeled samples for each class. A three-band false color image and the reference data are presented in Figure 2.16. Thirty samples for each class were randomly chosen from the reference data as training samples. The remaining samples composed the test set.

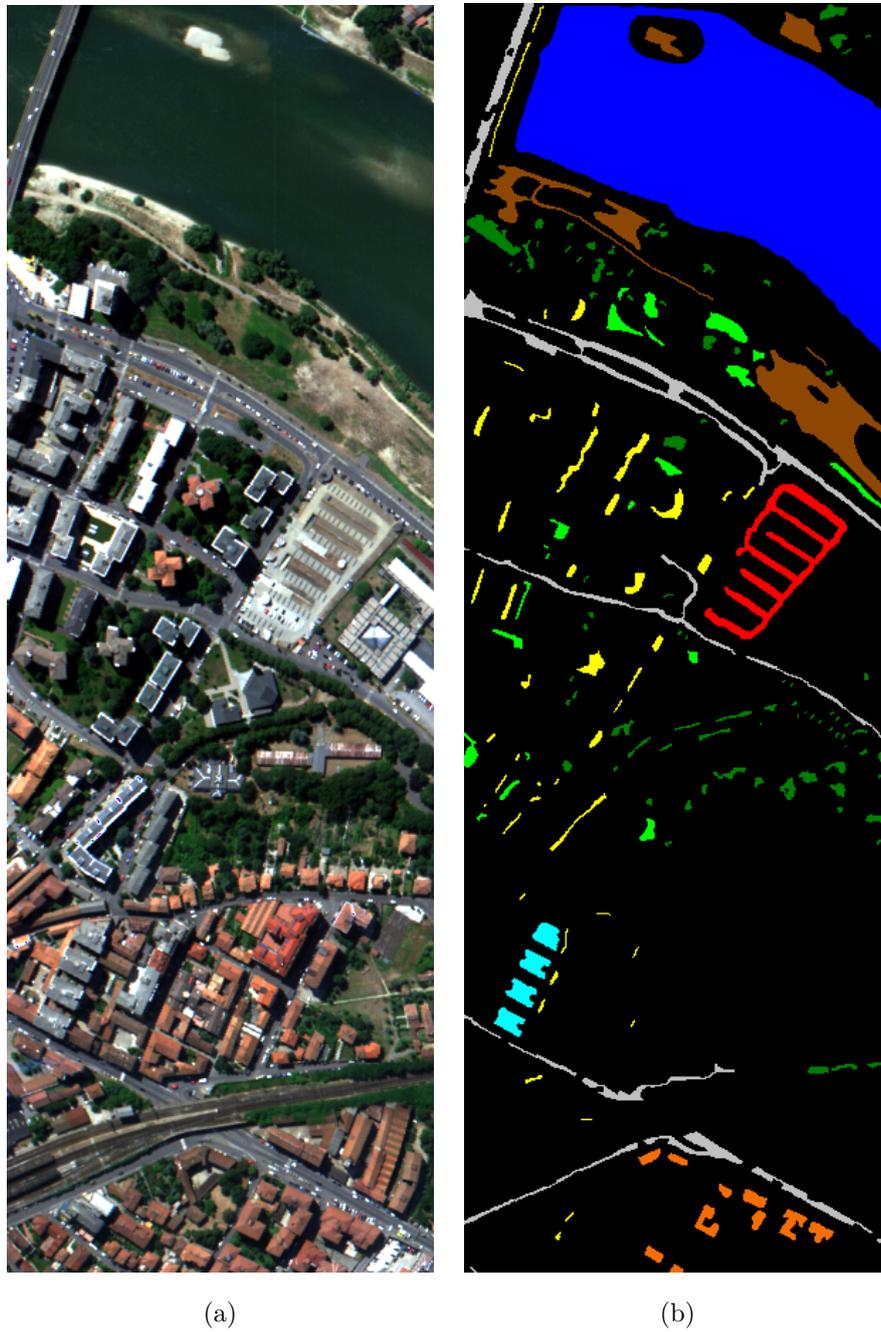


Figure 2.16: *Center of Pavia* image. (a) Three-band color composite (650, 558, and 478 nm). (b) Reference data: **Water**, **Trees**, **Meadows**, **Bricks**, **Bare soil**, **Asphalt**, **Bitumen**, **Tile**, and **Shadows**.

Table 2.2: Classification Accuracies in Percentage for the *Indian Pines* Image Using Pixelwise and Spectral-Spatial Approaches, along with Information Classes and Number of Labeled Samples (No. of Samp.): Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	No. of Samp.	Pixelwise methods			ECHO	Methods from Section 2.1	
		3-NN	ML	SVM		SVMMRF-NE	SVMMRF-E
OA	-	66.27	75.41	78.17	82.64	92.05	91.83
AA	-	76.77	79.61	85.97	83.75	95.83	95.69
κ	-	62.04	72.25	75.33	80.38	90.93	90.71
1 - Corn-no till	1434	41.84	71.39	78.18	83.45	93.28	98.48
2 - Corn-min till	834	62.24	63.01	69.64	75.13	83.93	90.82
3 - Corn	234	73.37	85.87	91.85	92.39	99.46	98.37
4 - Soybeans-no till	968	67.43	79.43	82.03	90.10	98.58	98.91
5 - Soybeans-min till	2468	53.91	52.65	58.95	64.14	82.09	76.92
6 - Soybeans-clean till	614	64.72	85.99	87.94	89.89	97.70	97.34
7 - Alfalfa	54	84.62	48.72	74.36	48.72	97.44	97.44
8 - Grass/pasture	497	86.35	93.51	92.17	94.18	97.54	97.54
9 - Grass/trees	747	91.97	94.69	91.68	96.27	97.70	97.56
10 - Grass/pasture-mowed	26	100	36.36	100	36.36	100	100
11 - Hay-windrowed	489	95.67	97.72	97.72	97.72	99.54	99.54
12 - Oats	20	80.00	100	100	100	100	100
13 - Wheat	212	99.38	98.15	98.77	98.15	99.38	99.38
14 - Woods	1294	86.17	95.42	93.01	94.21	98.39	99.04
15 - Bldg-Grass-Tree-Drives	380	45.15	73.03	61.52	81.52	88.18	79.70
16 - Stone-steel towers	95	95.56	97.78	97.78	97.78	100	100

2.4.2 Experimental results

2.4.2.1 Classification of the Indian Pines image

We compare the following methods for classification of the *Indian Pines* data set:

★ *Pixelwise methods:*

1. *3-NN*. First, a feature extraction on the original 200-band image \mathbf{X} was applied, using the method of PCFA [71] to get a 10-band image \mathbf{Y}_{IN} . Then, the 3-NN classification was performed [40], using the SAM distance measure.
2. *ML*. The ML technique was applied on the 10-band image \mathbf{Y}_{IN} feature vectors.
3. *SVM*. An SVM classification on the 200-band image was performed, using the multiclass one *versus* one SVM classifier with the Gaussian RBF kernel. The optimal parameters C and γ were chosen by fivefold cross validation: $C = 128$, $\gamma = 2^{-6}$.

★ *Previous spectral-spatial methods:*

4. *ECHO*. The ECHO classification was performed on the 10-band image \mathbf{Y}_{IN} , using the MultiSpec software [74, 81]. We have tested performances of the algorithm with different ranges of parameters, and here the best classification accuracies are reported.

Table 2.3: Classification Accuracies in Percentage for the *Indian Pines* Image Using Segmentation-Based Approaches: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	Methods from Section 2.2			Methods from Section 2.3						
	WH +MV	EM +MV	RHSEG +MV	SVM -WH	ML MSF	MLMSF +MV	SVM MSF	SVMMMSF +MV	MC- MSF	MSSC- MSF
OA	86.63	83.60	90.86	85.99	76.25	89.50	88.41	91.80	86.66	92.32
AA	91.61	85.34	93.96	86.95	82.32	91.45	91.57	94.28	92.58	94.22
κ	84.83	81.43	89.56	83.98	73.31	88.01	86.71	90.64	84.82	91.19
1	94.22	89.09	90.46	80.35	75.94	92.63	90.97	93.21	83.82	89.74
2	78.06	75.64	83.04	71.94	58.29	71.05	69.52	96.56	74.62	86.99
3	88.59	65.22	95.65	73.37	67.93	92.93	95.65	95.65	96.74	95.11
4	96.30	88.14	92.06	98.91	86.29	96.08	98.04	93.91	93.36	91.84
5	68.82	65.67	84.04	80.48	50.37	81.76	81.97	81.97	72.91	89.16
6	90.78	95.04	95.39	84.75	94.68	95.57	85.99	97.16	95.92	97.34
7	94.87	94.87	92.31	94.87	87.18	92.31	94.87	94.87	94.87	94.87
8	95.08	93.96	94.41	95.30	95.30	96.42	94.63	94.63	98.21	94.63
9	97.99	96.41	97.56	92.97	94.55	97.13	92.40	97.27	97.70	97.85
10	100	100	100	100	45.45	81.82	100	100	100	100
11	99.54	99.32	99.54	99.54	97.04	99.32	99.77	99.77	99.54	99.77
12	100	40.00	100	100	100	100	100	100	100	100
13	99.38	98.77	98.15	99.38	98.15	98.15	99.38	99.38	99.38	99.38
14	97.11	96.70	98.63	99.36	96.70	97.19	97.59	99.68	98.47	99.44
15	69.39	66.67	82.12	55.45	71.52	73.03	68.79	68.79	77.88	73.64
16	95.56	100	100	64.44	97.78	97.78	95.56	95.56	97.78	97.78

- ★ **Methods from Section 2.1.** Context classification using closest fixed neighborhoods, SVM and MRFs was performed, using two different approaches discussed in Section 2.1 and Paper 1:
- 5. *SVMMRF-NE*. The optimal choice of parameters is discussed in Paper 1. Here, $\beta = 1$ is chosen.
- 6. *SVMMRF-E*. The parameter $\beta = 2$ is chosen.

- ★ **Methods from Section 2.2.** The segmentation of the considered image was performed, using watershed (WH), EM and RHSEG algorithms. Then, the results of the pixelwise SVM classification were combined with the segmentation results, using the majority voting approach, as discussed in Section 2.2 and Papers 2, 3, and 6:
- 7. *WH+MV*. The extension of a watershed to the case of hyperspectral images is investigated in Paper 2, where different approaches are compared. Here, the one-band RCMG on the considered image was computed, followed by the standard Vincent and Soille watershed algorithm [149].
- 8. *EM+MV*. We report the results of partitional clustering-based segmentation using the EM algorithm (classification results using the ISODATA technique can be found in Paper 3). For the EM algorithm, the upper bound on the number of clusters was chosen equal to 17 (typically slightly superior to the number of classes).
- 9. *RHSEG+MV*. Since some classes have very similar spectral responses in the considered image (for instance, three classes of corn and three classes of soybeans), we have set $S_{weight} = 0.0$ for the RHSEG method. A segmentation map at the relevant level of hierarchy was chosen with the program HSEGVier.

★ **Methods from Section 2.3:**

10. *SVM-WH*. Marker selection using probabilistic SVM (Section 2.3.1.2, the choice of parameters is discussed in Paper 5) followed by marker-controlled watershed (Section 2.3.2.1) was performed.
11. *MLMSF*. Marker selection using ML discriminant values (Section 2.3.1.1) followed by construction of an MSF, using the SAM dissimilarity measure (the following methods were also applied using the SAM dissimilarity measure; Section 2.3.2.2) was applied. The optimal value of the parameter $t = 0.001$ was experimentally derived.
12. *MLMSF+MV*. MLMSF, followed by the optional majority voting within connected components step (see Section 2.3.2.2) was performed.
13. *SVMMMSF*. Marker selection using probabilistic SVM (Section 2.3.1.2), followed by construction of an MSF (Section 2.3.2.2) was applied.
14. *SVMMMSF+MV*. SVMMMSF, followed by the optional majority voting within connected components step (see Section 2.3.2.2) was performed.
15. *MC-MSF*. MC marker selection using 3-NN, ML, and SVM classification maps (Section 2.3.1.3), followed by construction of an MSF (Section 2.3.2.2) was applied.
16. *MSSC-MSF*. MSSC marker selection using WH+MV, EM+MV, and RHSEG+MV classification maps (Section 2.3.1.3), followed by construction of an MSF (Section 2.3.2.2) was performed.

Tables 2.2 and 2.3 report global (OA, AA, and kappa coefficient) and class-specific classification accuracies for the described methods. The corresponding classification maps are depicted in Figure 2.17. From the tables and from the figure, the following **conclusions** can be drawn:

Performances of pixelwise methods: The *SVM* method gives the highest accuracies among all the pixelwise classification techniques.

Pixelwise versus spectral-spatial methods: All the spectral-spatial approaches yield higher classification accuracies (except the *MLMSF* method, which performs slightly worse as compared to the *SVM* method) and classification maps with more homogeneous regions when compared to pixelwise methods.

Proposed spectral-spatial methods versus the standard ECHO context classifier: All the proposed spectral-spatial approaches significantly outperform the standard ECHO context classifier (except the *MLMSF* method).

Methods from Section 2.1: The *SVMMRF-NE* and *SVMMRF-E* accuracies are not significantly different (following the results of McNemar’s test, using 5% level of significance). The *SVMMRF-NE*

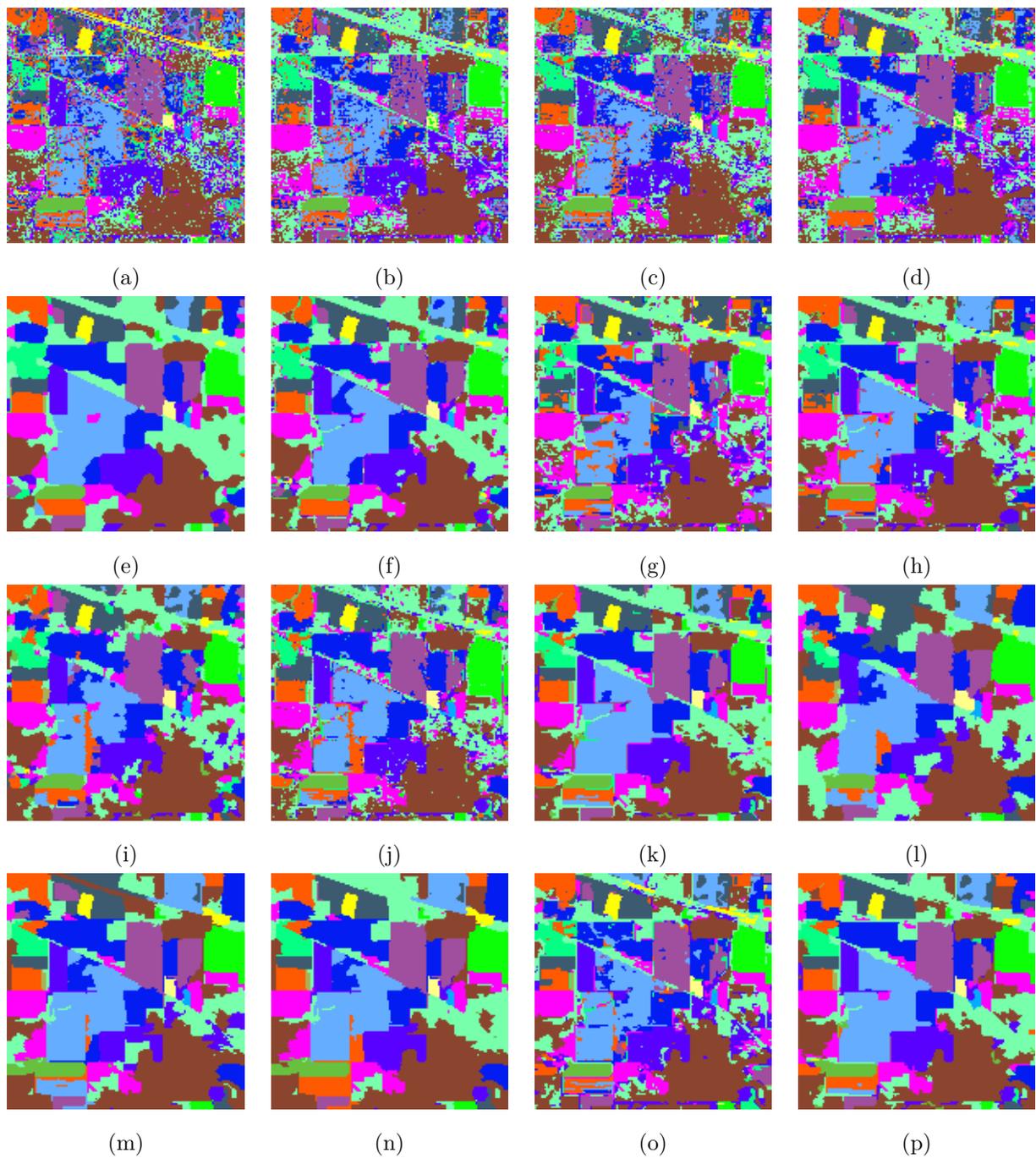


Figure 2.17: Classification maps for the *Indian Pines* image. (a) 3-NN. (b) ML. (c) SVM. (d) ECHO. (e) SVMRF-NE. (f) SVMRF-E. (g) MLMSF. (h) MLMSF+MV. (i) WH+MV. (j) EM+MV. (k) RHSEG+MV. (l) SVM-WH. (m) SVMMSF. (n) SVMMSF+MV. (o) MC-MSF. (p) MSSC-MSF.

Table 2.4: Classification Accuracies in Percentage for the *Hekla* Image Using Pixelwise and Spectral-Spatial Approaches, along with Information Classes and Number of Labeled Samples (No. of Samp.): Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	No. of Samp.	Pixelwise methods			ECHO	Methods from Section 2.1	
		3-NN	ML	SVM		SVMMRF-NE	SVMMRF-E
OA	-	90.17	96.18	88.56	96.63	98.90	98.63
AA	-	86.60	96.99	89.44	97.67	99.62	99.23
κ	-	88.64	95.59	86.91	96.12	98.73	98.42
Andesite lava 1970	342	87.67	98.97	88.36	99.66	100	100
Andesite lava 1980 I	708	93.02	98.94	87.25	99.24	99.85	100
Andesite lava 1980 II	1496	92.39	94.26	88.24	94.26	100	100
Andesite lava 1991 I	2739	97.10	94.01	84.94	94.38	96.24	96.47
Andesite lava 1991 II	410	68.33	96.39	93.33	96.94	100	100
Andesite lava with moss cover	1023	91.98	98.15	94.24	98.25	100	98.97
Hyaloclastite formation	684	87.07	98.74	87.54	99.37	100	100
Lava with tephra and scoria	700	82.00	96.15	91.69	96.31	99.38	99.38
Rhyolite	404	86.44	92.37	85.88	95.48	100	100
Scoria	550	55.00	97.60	74.20	99.60	100	97.80
Firn and glacier ice	458	100	100	100	100	100	100
Snow	713	98.19	98.34	97.59	98.49	100	98.19

method tends to smooth region borders, while the *SVMMRF-E* approach results in a classification map, where borders between regions are well preserved.

Methods from Section 2.2: The RHSEG segmentation technique leads to the best classification accuracies, when compared to other segmentation approaches. The spectral-based EM approach performs worse than the spatial-based watershed technique for segmentation of the Indian Pines data. Consequently, the spatial-based segmentation techniques appear to be preferable when an image contains classes with similar spectral responses.

Methods from Section 2.3: The best overall classification accuracy is obtained when using an MSSC approach for marker selection followed by construction of an MSF (the *MSSC-MSF* method, which significantly outperforms the *MC-MSF* method). The best AA is obtained when applying the *SVMMRF+MV* technique. The optional majority voting within connected components step (applied after construction of an MSF) additionally improved classification accuracies. The ML-based marker selection approach performs worse as compared to two other marker selection methods. The MSF-based method for marker-controlled region growing significantly outperforms the marker-controlled watershed technique. Furthermore, the *SVMMRF+MV* method provides at its output an accurate segmentation map, where almost no oversegmentation is present.

Methods yielding the most accurate classification results: The best classification results are obtained when applying the *MSSC-MSF*, *SVMMRF-NE*, *SVMMRF-E*, and *SVMMRF+MV* methods. For all these methods, the classification accuracies are not significantly different (following the results of McNemar’s test, using 5% level of significance), and the corresponding classification maps look very similar. The *RHSEG+MV* technique from Section 2.2 yields slightly lower accuracies, although the classification map obtained by this method is comparable to the maps obtained by the four winning methods.

Table 2.5: Classification Accuracies in Percentage for the *Hekla* Image Using Segmentation-Based Approaches: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	Methods from Section 2.2			Methods from Section 2.3			
	WH+MV	EM+MV	RHSEG +MV	SVMMSF	SVMMSF +MV	MC- MSF	MSSC- MSF
OA	96.93	87.85	94.72	90.34	98.96	99.08	98.41
AA	97.57	82.78	95.03	94.89	98.45	99.07	97.52
κ	96.45	85.85	93.94	89.04	98.80	98.93	98.16
Andesite lava 1970	97.26	96.23	97.26	100	100	100	98.29
Andesite lava 1980 I	97.72	99.39	99.85	92.11	100	99.85	100
Andesite lava 1980 II	95.85	98.27	100	96.96	99.86	99.24	100
Andesite lava 1991 I	95.28	99.44	88.69	73.19	99.55	99.37	98.25
Andesite lava 1991 II	99.72	78.33	68.89	88.89	88.89	100	77.78
Andesite lava with moss cover	100	96.81	98.15	98.46	98.46	98.66	100
Hyaloclastite formation	97.16	2.37	99.53	99.53	99.68	98.74	100
Lava with tephra and scoria	97.54	99.54	99.69	95.08	97.38	98.77	100
Rhyolite	98.87	96.05	98.02	96.89	100	99.72	99.72
Scoria	93.80	29.80	93.20	97.60	97.60	96.00	96.60
Firn and glacier ice	100	100	100	100	100	100	100
Snow	97.59	97.13	97.13	100	100	98.49	99.55

2.4.2.2 Classification of the Hekla image

We compare the methods described in the previous subsection for classification of the *Hekla* image (except the *SVM-WH*, *MLMSF*, and *MLMSF+MV* methods, since they have been less robust and less successful in classifying the previous data set, when compared to other similar techniques). Tables 2.4 and 2.5 give global and class-specific accuracies for these methods (the results corresponding to the optimal parameters are reported). Most of the corresponding classification maps are shown in Figure 2.18. From these results, the following conclusions can be derived:

Performances of pixelwise methods: The *ML* method yields the highest accuracies among all the pixelwise classification techniques. However, since the test set is small when compared to the image dimensions, it is important to visually compare classification maps. When comparing the three pixelwise classification maps with the *MC-MSF* and *SVMMRF-NE* maps (yielding the highest global classification accuracies), it can be concluded that the *SVM* map is the most similar one to these two spectral-spatial approaches.

Pixelwise versus spectral-spatial methods: Spectral-spatial approaches typically outperform pixelwise classification methods. Although *EM+MV*, *RHSEG+MV*, and *SVMMSF* classification accuracies are lower when compared to the *ML* accuracies, the similar visual comparison of classification maps yields the conclusion about better performances of the spectral-spatial techniques.

Proposed spectral-spatial methods versus the standard ECHO context classifier: From the visual comparison of the classification maps, it can be concluded that the proposed spectral-spatial methods are preferable to the standard *ECHO* classifier.

Methods from Section 2.1: The *SVMMRF-NE* method yields non-significantly higher accuracies

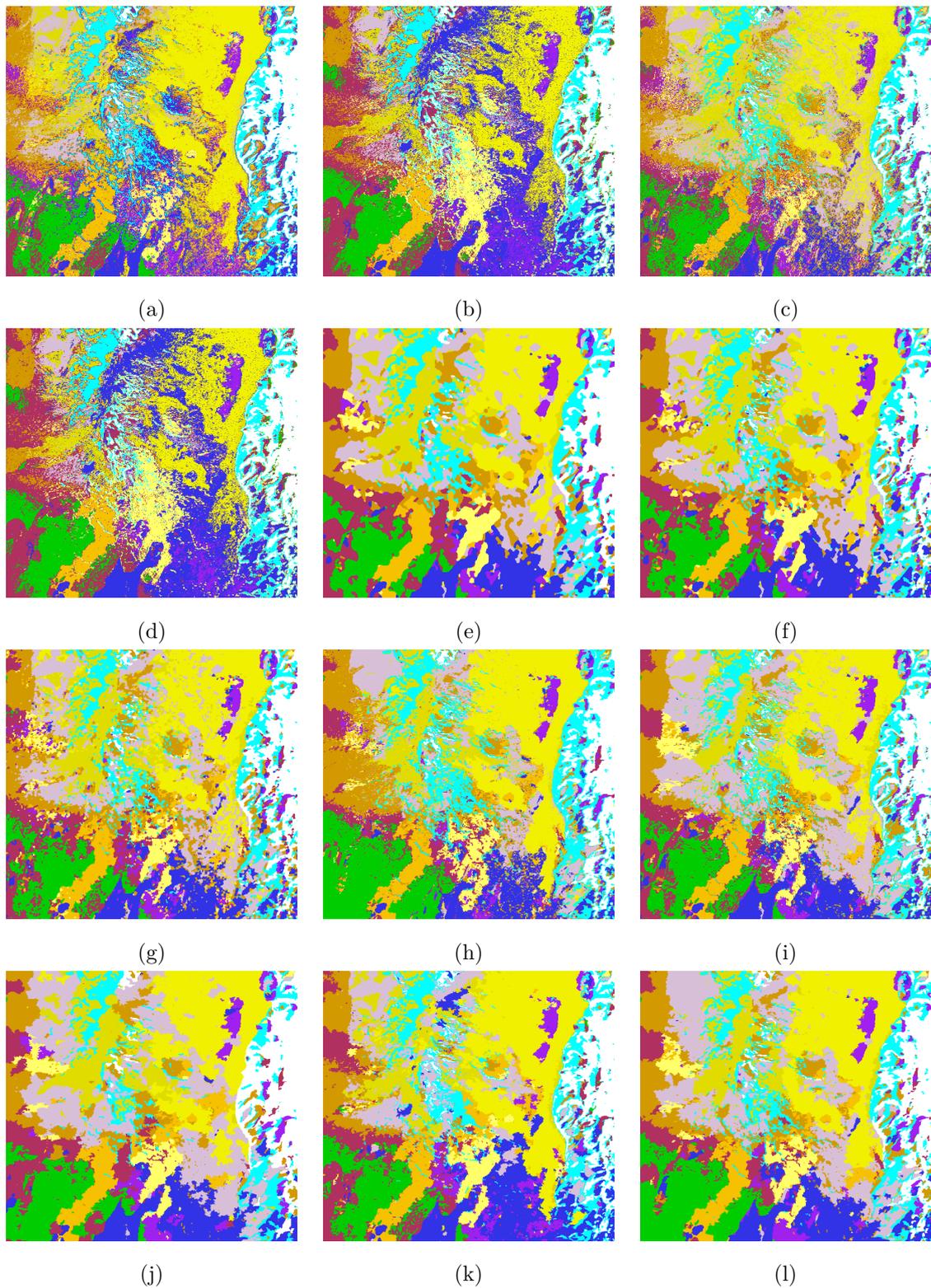


Figure 2.18: Classification maps for the *Hekla* image. (a) 3-NN. (b) ML. (c) SVM. (d) ECHO. (e) SVMMRF-NE. (f) SVMMRF-E. (g) WH+MV. (h) EM+MV. (i) RHSEG+MV. (j) SVMMSF+MV. (k) MC-MSF. (l) MSSC-MSF.

when compared to the *SVMMRF-E* approach. The *SVMMRF-E* classification map contains several small regions present in other spectral-spatial classification maps and missed in the *SVMMRF-NE* map.

Methods from Section 2.2: The watershed segmentation technique leads to the best classification accuracies, when compared to other segmentation approaches. The same conclusion as for the previous data set can be drawn that the spatial-based segmentation approaches are preferable to the spectral-based techniques when an image contains classes with similar spectral responses.

Methods from Section 2.3: The best global accuracies are obtained when applying the *MC-MSF* method. The *SVMMRF+MV* and *MSSC-MSF* techniques yield non-significantly lower accuracies. From the visual comparison of all the classification maps, the *MSSC-MSF* appears to be the most accurate one.

Methods yielding the most accurate classification results: The *MC-MSF* and *SVMMRF-NE* methods yield the highest OA and AA, respectively. Although the *SVMMRF-E*, *SVMMRF+MV*, and *MSSC-MSF* approaches give slightly lower accuracies, the corresponding classification maps are comparable (and may be even more accurate) when compared to the two winning techniques.

2.4.2.3 Classification of the urban data sets

Finally, we compare the considered methods for classification of the urban data sets (the *University of Pavia* and the *Center of Pavia* images, which we call in the following the *University* and the *Center* images, respectively). Tables 2.6-2.9 report global and class-specific accuracies for the *University* and the *Center* images, respectively (the results obtained with the optimal parameters are reported). Some of the corresponding classification maps are depicted in Figures 2.19-2.22. We have included in Table 2.6 accuracies of mathematical morphology-based classification using SVM, principal components and Extended Morphological Profiles (*EMP*). This method is especially suitable for analyzing urban images. The classification results are taken from Plaza *et al.* [111], where the same training and test samples were used for classification. Other results of spectral-spatial classification of the Pavia data sets can be found in [4, 43, 45]. From the tables and the figures, the following conclusions can be made:

Performances of pixelwise methods: The *SVM* method gives the highest accuracies for the *University* image and almost the highest accuracies for the *Center* image (the ML technique non-significantly outperforms the SVM method for this data set).

Pixelwise versus spectral-spatial methods: All the spectral-spatial approaches yield significantly higher classification accuracies and classification maps with more homogeneous regions, as compared to pixelwise methods.

Proposed spectral-spatial methods versus previous context classifiers: The proposed spectral-spatial approaches yield better classification results when compared to the previous context classifiers (except for the *University* image, the *SVMMRF-NE*, *WH+MV*, and *SVMMRF* methods give lower accuracies as compared to the ECHO approach).

Methods from Section 2.1: For both data sets, the *SVMMRF-E* method outperforms the *SVMMRF-NE* technique in terms of accuracies. This advantage of the edge-based *SVMMRF-E* method for classification of urban images can be explained as follows: Images of urban areas contain small and complex spatial structures, such as shadows and trees. Therefore, the inclusion of the edge information into the context-based regularization improves classification performances.

Methods from Section 2.2: The EM and the RHSEG segmentation methods lead to the best OA and AA, respectively. Thus, if an image contains classes with dissimilar spectral responses, cluster-based segmentation techniques yield good results. Furthermore, spectral-based methods can be preferable if an image contains a lot of small and complex spatial structures that risk being assimilated with larger neighboring structures when performing segmentation.

Table 2.6: Classification Accuracies in Percentage for the *University of Pavia* Image Using Pixelwise and Spectral-Spatial Approaches, along with Information Classes and Training-Test Samples: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	Samples		Pixelwise methods			Previous methods		Methods from Section 2.1	
	Train	Test	3-NN	ML	SVM	ECHO	EMP	SVMMRF-NE	SVMMRF-E
OA	-	-	68.38	79.06	81.01	87.58	85.22	86.89	87.63
AA	-	-	77.21	84.85	88.25	92.16	90.76	92.12	93.41
κ	-	-	59.85	72.90	75.86	83.90	80.86	83.14	84.07
Asphalt	548	6304	64.96	76.43	84.93	87.98	95.36	97.32	96.88
Meadows	540	18146	63.18	75.99	70.79	81.64	80.33	76.59	77.43
Gravel	392	1815	62.31	64.57	67.16	76.91	87.61	66.34	72.34
Trees	524	2912	95.95	97.08	97.77	99.31	98.37	99.31	99.28
Metal sheets	265	1113	99.73	99.91	99.46	99.91	99.48	100	99.91
Bare soil	532	4572	57.42	70.03	92.83	93.96	63.72	98.25	98.12
Bitumen	375	981	82.67	90.62	90.42	92.97	98.87	95.11	97.35
Bricks	514	3364	77.08	90.10	92.78	97.35	95.41	98.57	99.46
Shadows	231	795	91.57	98.87	98.11	99.37	97.68	97.61	99.87

Table 2.7: Classification Accuracies in Percentage for the *University of Pavia* Image Using Segmentation-Based Approaches: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	Methods from Section 2.2			Methods from Section 2.3			
	WH+MV	EM+MV	RHSEG +MV	SVMMSF	SVMMSF +MV	MC- MSF	MSSC- MSF
OA	85.42	94.00	93.85	84.14	91.08	87.98	97.90
AA	91.31	93.13	97.07	92.35	94.76	92.05	98.59
κ	81.30	91.93	91.89	79.71	88.30	84.32	97.18
Asphalt	93.64	90.10	94.77	93.05	93.16	87.01	98.00
Meadows	75.09	95.99	89.32	72.30	85.65	83.24	96.67
Gravel	66.12	82.26	96.14	89.15	89.15	75.37	97.80
Trees	98.56	85.54	98.08	87.02	91.24	98.97	98.83
Metal sheets	99.91	100	99.82	99.91	99.91	99.91	99.91
Bare soil	97.35	96.72	99.76	97.11	99.91	93.24	100
Bitumen	96.23	91.85	100	98.57	98.57	95.11	99.90
Bricks	97.92	98.34	99.29	95.66	99.05	97.00	99.76
Shadows	96.98	97.36	96.48	98.36	96.23	98.62	96.48

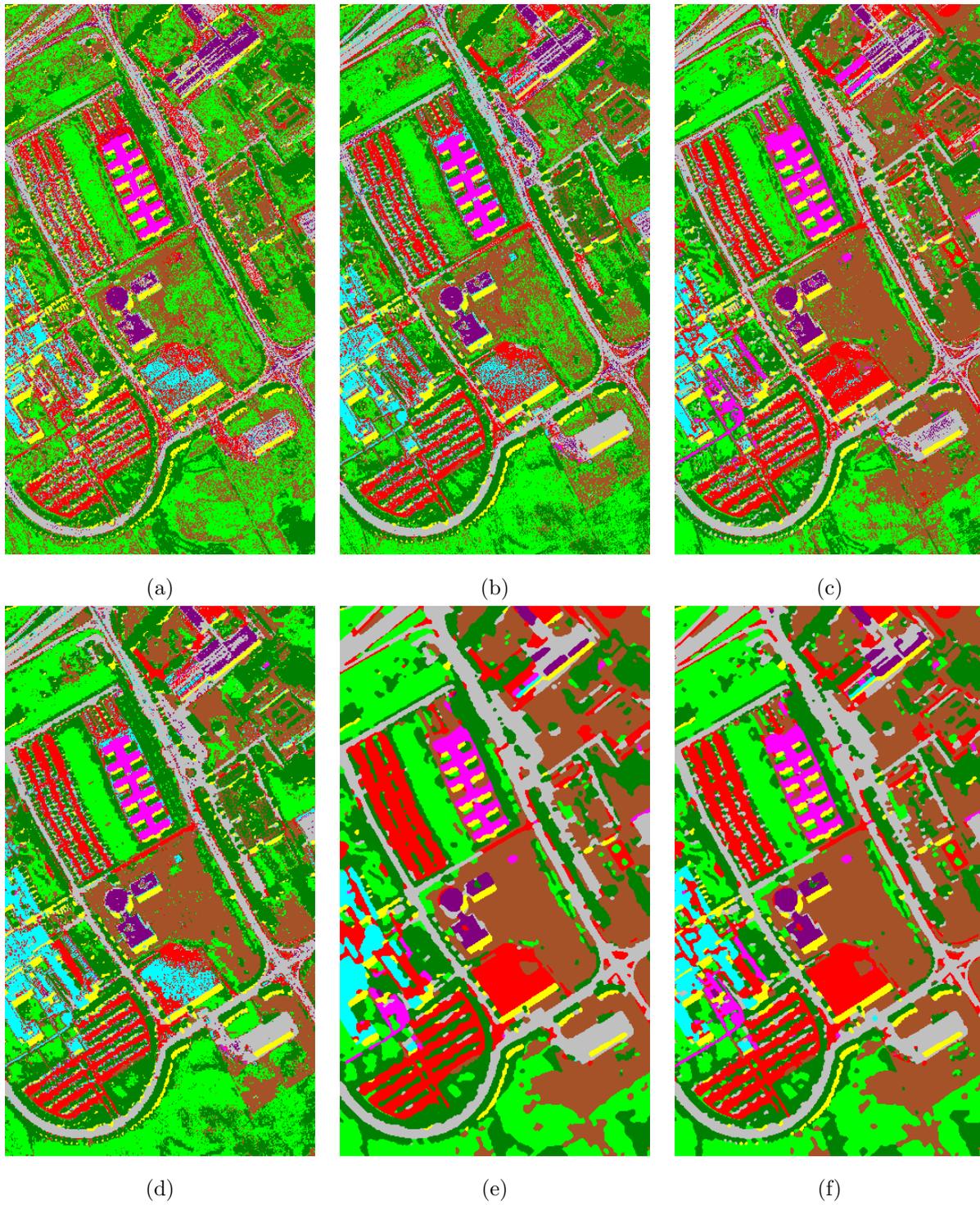


Figure 2.19: Classification maps for the *University of Pavia* image. (a) 3-NN. (b) ML. (c) SVM. (d) ECHO. (e) SVMRF-NE. (f) SVMRF-E.

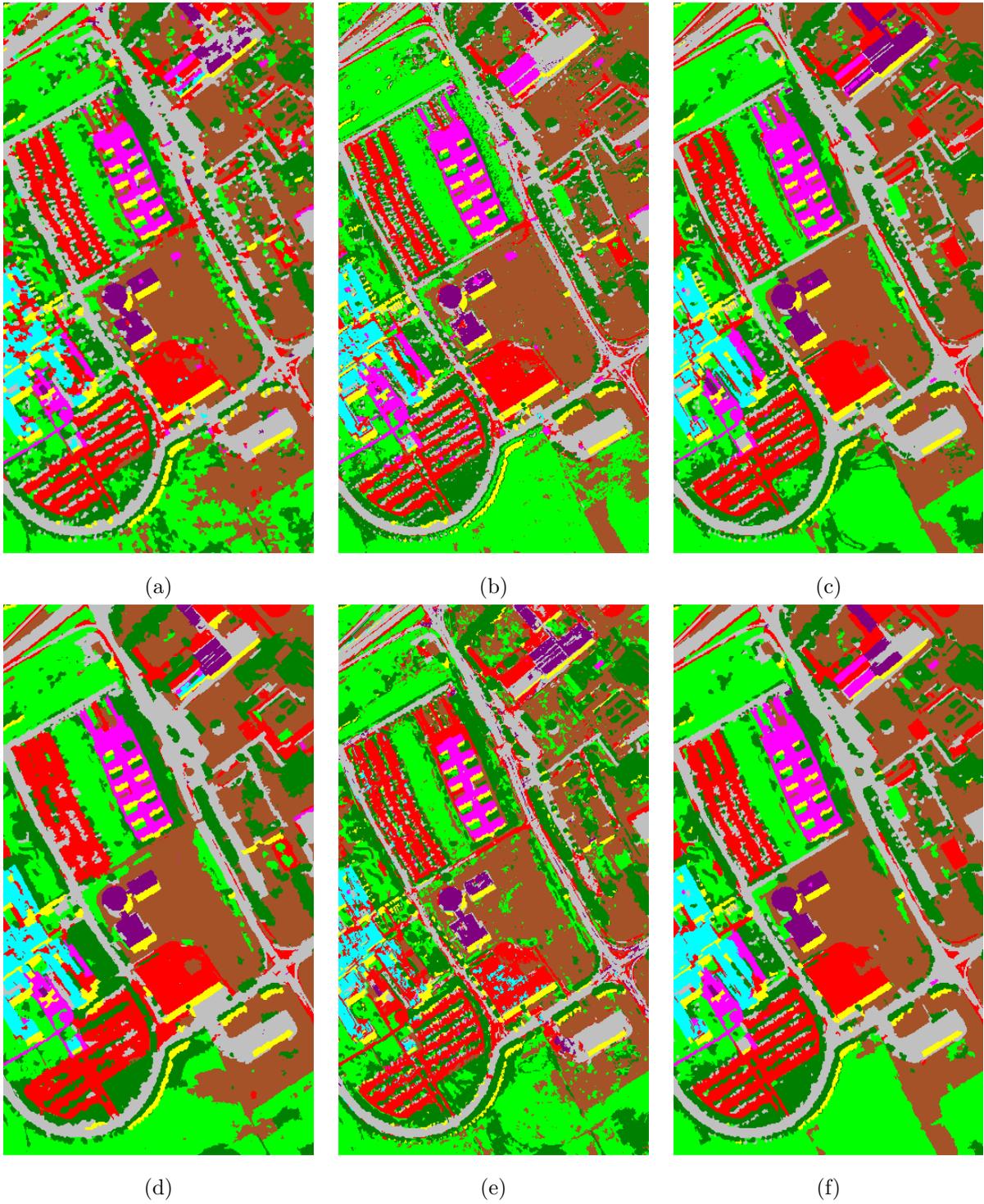


Figure 2.20: Classification maps for the *University of Pavia* image. (a) WH+MV. (b) EM+MV. (c) RHSEG+MV. (d) SVMMSF+MV. (e) MC-MSF. (f) MSSC-MSF.

Table 2.8: Classification Accuracies in Percentage for the *Center of Pavia* Image Using Pixelwise and Spectral-Spatial Approaches, along with Information Classes and Number of Labeled Samples (No. of Samp.): Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	No. of Samp.	Pixelwise methods			ECHO	Methods from Section 2.1	
		3-NN	ML	SVM		SVMMRF-NE	SVMMRF-E
OA	-	90.30	95.75	95.64	96.22	97.38	97.60
AA	-	80.51	91.13	90.60	92.47	94.13	94.59
κ	-	83.90	92.91	92.71	93.70	95.61	95.97
Water	34352	98.87	99.35	99.78	99.35	100	100
Trees	2627	70.54	92.72	90.26	93.88	96.03	96.03
Meadows	1788	90.16	87.14	96.42	88.17	99.66	100
Bricks	2140	74.55	79.34	64.03	82.56	62.89	64.17
Bare soil	5365	79.16	90.03	88.38	91.13	94.64	95.11
Asphalt	5568	68.04	89.67	90.45	90.45	95.03	96.01
Bitumen	972	60.62	87.90	87.47	92.36	98.94	100
Tile	1112	93.16	95.66	98.71	95.93	100	100
Shadows	2146	89.46	98.39	99.95	98.39	100	100

Table 2.9: Classification Accuracies in Percentage for the *Center of Pavia* Image Using Segmentation-Based Approaches: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	Methods from Section 2.2			Methods from Section 2.3			
	WH+MV	EM+MV	RHSEG+MV	SVMMSF	SVMMSF+MV	MC-MSF	MSSC-MSF
OA	96.26	97.07	97.06	96.37	96.62	97.04	97.78
AA	92.08	92.00	93.90	92.55	92.78	94.34	94.82
κ	93.75	95.08	95.08	93.93	94.35	95.04	96.28
Water	99.81	99.85	99.80	99.87	99.87	100	99.81
Trees	92.03	94.49	92.26	88.83	90.33	87.22	94.99
Meadows	97.78	81.40	97.04	99.49	96.93	97.38	94.08
Bricks	64.22	73.22	72.46	60.52	61.09	87.30	75.55
Bare soil	90.68	94.47	90.98	90.42	92.95	89.39	95.39
Asphalt	91.93	99.31	96.14	93.93	93.93	93.28	98.39
Bitumen	93.31	97.24	98.51	99.89	99.89	96.28	99.58
Tile	98.98	93.81	98.24	100	100	99.54	99.08
Shadows	100	94.23	99.67	100	100	98.68	96.50

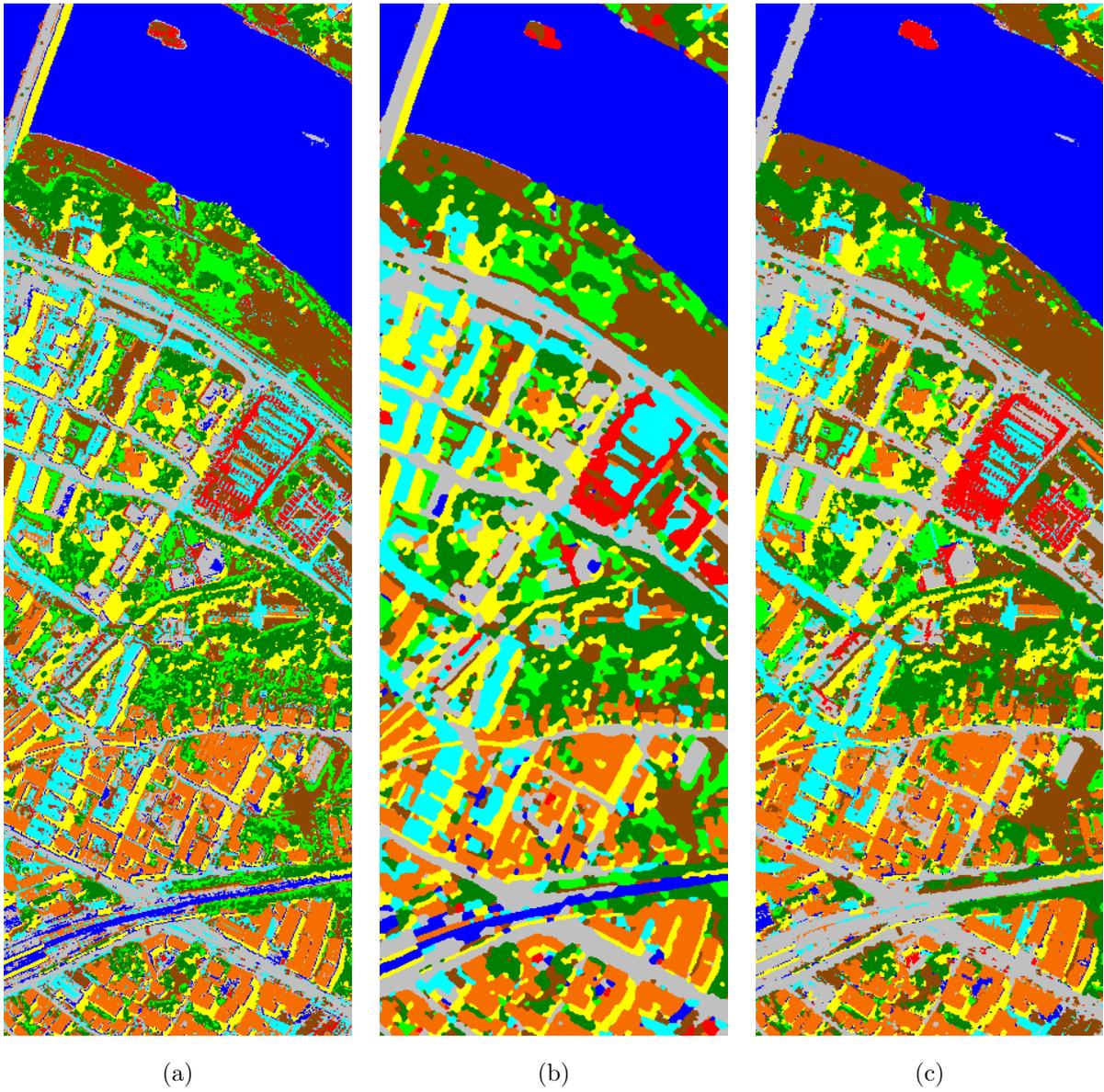


Figure 2.21: Classification maps for the *Center of Pavia* image. (a) SVM. (b) SVMMRF-E. (c) EM+MV.

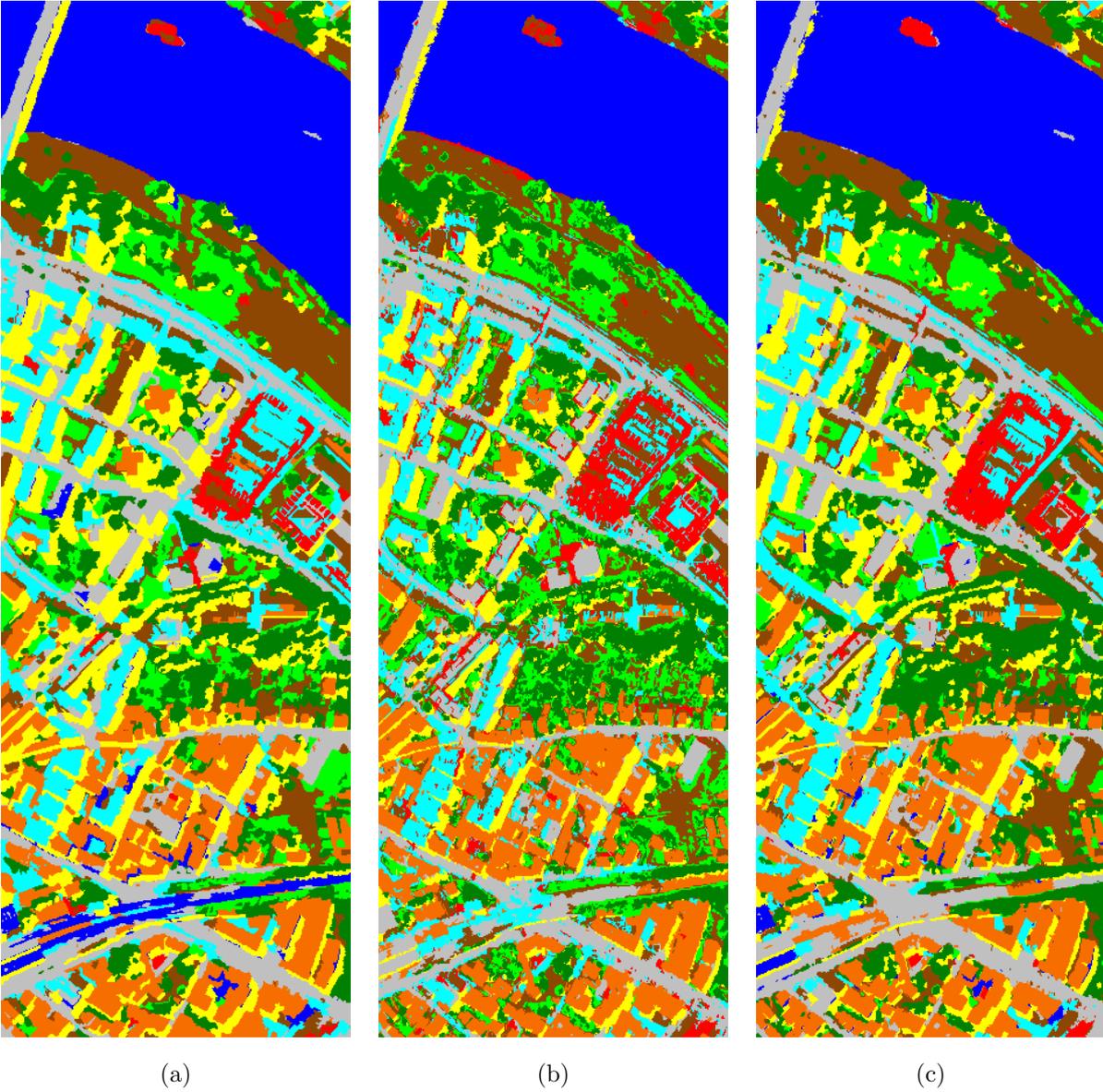


Figure 2.22: Classification maps for the *Center of Pavia* image. (a) SVMMSF+MV. (b) MC-MSF. (c) MSSC-MSF.

Table 2.10: Methods Yielding the Most Accurate Classification Results (😊) for the Four Considered Data Sets.

Method	Indian Pines	Hekla	University	Center
SVMMRF-NE	😊	😊		😊
SVMMRF-E	😊	😊		😊
EM+MV			😊	
RHSEG+MV			😊	
SVMMSF+MV	😊	😊	😊	
MC-MSF		😊		
MSSC-MSF	😊	😊	😊	😊

Methods from Section 2.3: The best global classification accuracies and the most homogeneous classification maps are obtained when applying the *MSSC-MSF* technique. The *MC-MSF* and *SVMMSF+MV* methods also perform well, although they yield less accurate results than the *MSSC-MSF* method.

Methods yielding the most accurate classification results: For both data sets, the best classification results are obtained when the *MSSC-MSF* method was applied (more than 97% of test samples are correctly classified). The *EM+MV*, *RHSEG+MV* and *SVMMSF+MV* techniques also yield very accurate results. The *SVMMRF-NE* and *SVMMRF-E* methods yield close to the highest accuracies for the *Center* image, while for the *University* image they are less efficient when compared to methods from Sections 2.2 and 2.3. These results can be explained as follows: The MRF-based regularization includes the spatial context information from only the closest neighborhoods (in our case, eight-neighborhoods) when classifying an image. Therefore, this method is efficient only in the case if there is no large misclassified region in the initial pixelwise classification map (this assumption often holds). If such a region exists, the MRF-based method cannot reconstitute its true class label. This happens in the case of classification of the *University* image, where some relatively large regions remain misclassified. Thus, the segmentation-based context classification techniques appear to be robust procedures for classifying large regions with uncertain spectral properties.

2.4.3 Concluding discussion

Experimental evaluation of the proposed approaches for spectral-spatial classification of hyperspectral images has been conducted. Table 2.10 summarizes the methods yielding the most accurate classification results for the four considered images. The experimental results presented in this section lead to the following conclusions:

1. The SVM method has proven to be efficient for pixelwise classification of high-dimensional data.
2. It is advantageous to consider spatial dependencies between pixels when performing classification.
3. The spectral-spatial classification methods developed in the frame of this thesis significantly outperform in terms of accuracies the previously proposed *ECHO* and *EMP* techniques.
4. The MRF-based regularization has proven to be a powerful tool for contextual image analysis. The developed MRF-based classification approaches (the *SVMMRF-NE* and *SVMMRF-E* methods) are efficient and sufficiently robust for classifying different kinds of images. The *SVMMRF-E* technique includes the edge information into the context-based regularization yielding classification maps with accurate borders. Hence, we recommend this technique for hyperspectral image classification, especially for images which do not contain large regions with uncertain spectral properties. If such regions exist, there is a risk to have large misclassified regions in the pixelwise classification map, and the MRF-

based method cannot reconstitute their true class label. In this case, the segmentation-based context classification techniques are preferable.

5. Based on the analysis of several unsupervised segmentation approaches, we conclude that the spatial-based segmentation techniques appear to be preferable when an image contains classes with similar spectral responses (e.g., images of a vegetated area, when an accurate distinction between different species of vegetation is required). If an image contains classes with dissimilar spectral responses, cluster-based segmentation techniques yield good results. Furthermore, cluster-based methods can be helpful if an image contains a lot of small and complex structures that risk being assimilated with larger adjacent regions (for instance, images of an urban area). Finally, the segmentation methods working both in the spatial and the spectral domain (for instance, the RHSEG method) are promising in obtaining accurate segmentation results on different images.

6. The proposed method of combining a pixelwise SVM classification map and a segmentation map, using a majority vote rule, has proven to be an efficient context classification technique.

7. Region growing segmentation based on automatically derived markers is a promising technique for contextual image analysis, since it allows one to mitigate dependence of segmentation performances on the selected homogeneity criteria. The use of probabilistic classification results for selecting the most reliable pixels as region markers has proven to be an efficient approach. In particular, it is advantageous to use an SVM classifier, spatial information and MC approaches in the marker selection procedure. The MSF-based method for marker-controlled region growing significantly outperforms the marker-controlled watershed technique, and thus has proven to be an efficient and robust segmentation technique. The *MSSC-MSF* method has yielded the best or close to the best classification results for all the considered data sets and thus is highly recommended for accurate context classification of hyperspectral data. Another recommended choice is the *SVMMSF+MV* classifier, which also yields high accuracies and is easier to implement when compared to the *MSSC-MSF* method.

2.5 Implementation Issues

This section addresses the issue of computational loads when processing hyperspectral data. First, we briefly compare the computational costs of the proposed spectral-spatial classification approaches. Then, we discuss possibilities of speeding up computational processes using parallel computing on commodity GPUs.

2.5.1 Consideration of the computational cost for the proposed classification methods

While spectral-spatial classification typically yields more accurate results when compared to pixelwise classification, including of the spatial information in a classifier often increases the computational cost of the classification method. We aimed to develop spectral-spatial classification approaches which are efficient both in terms of classification accuracies and computational complexity. Thus, when comparing the results of several classifiers, an important issue is the computational cost of each classifier.

Papers 2 and 5 discuss computational complexity of the methods from Sections 2.2 and 2.3, respectively. Here, we compare the computational costs of the proposed approaches by summarizing in Table 2.11 the processing time for classification of the *University of Pavia* image, using different methods. We conducted experiments on an Intel Core 2 Duo 2.40 GHz processor with 3.5 GB RAM. Note that for the *WH+MV*, *EM+MV*, *RHSEG+MV*, and *MSSC-MSF* techniques, pixelwise classification and unsupervised segmentation of an image can be executed at the same time on different processing units (since none of these tasks depends on the results of another task). Consequently, when we compute the approximate total time for these techniques, we assume that these tasks are performed at the same time. From these results, the following conclusions can be drawn:

Although the pixelwise SVM classifier gives good classification accuracies, it is a computationally demanding algorithm for a high-dimensional data and/or when the number of training samples is large [33,

Table 2.11: Processing Time in Seconds for Classification of the *University of Pavia* Image (“classif.” means classification)

Algorithm	Pixelwise classif. (training+classif.)	Unsupervised segmentation	Spatial regularization	Approximate total time
Pixelwise SVM	3240 + 99	-	-	3339
SVMMRF-NE	3240 + 106	-	98 (MRF regularization)	3444
SVMMRF-E	3240 + 106	-	104 (MRF regularization)	3450
WH+MV	3240 + 99	13	< 1 (majority voting)	3340
EM+MV	3240 + 99	3	< 1 (majority voting)	3340
RHSEG+MV	3240 + 99	118 ($S_{weight} = 0.0$) 885 ($S_{weight} = 0.1$)	< 1 (majority voting)	3340
SVMMSF+MV	3240 + 106	-	2 (marker selection) + 2 (MSF) + < 1 (majority voting)	3351
MSSC-MSF	3240 + 99	13 (WH) 3 (EM) 885 (RHSEG)	< 1 (majority voting) + 1 (marker selection) + 2 (MSF)	3343

45]. The training part of the SVM classification is the most time-consuming, in particular the tuning of parameters by cross validation. The processing time for the training part by means of the LIBSVM library was 3240 seconds, while classification part without and with the computation of probability estimates was executed in 99 and 106 seconds, respectively. However, in recent works, methods and parallel implementations to speed up the SVM training and classification have been proposed [23, 57, 62, 111].

The pixelwise SVM classification is a part of all the considered spectral-spatial classification approaches (leading to the highest accuracies). The other, context-related parts (unsupervised segmentation and spatial regularization) of the proposed approaches are much less time-consuming than the SVM classification algorithms. Thus, the approximate total time of the proposed classification techniques increases by only a few seconds (the longest *SVMMRF-E* method is executed in 3450 seconds, what is less than two minutes more than the processing time of the pixelwise SVM classification).

Another interesting observation can be made: Based on experimental evaluation of the proposed techniques, we have recommended to use *MSSC-MSF*, then *SVMMSF+MV* methods, which have proved to be the most efficient and robust for a variety of images. From Table 2.11, it can be seen that the *MSSC-MSF* method, yielding the highest classification accuracies, is also one of the fastest spectral-spatial classification techniques. The *SVMMSF+MV* approach is only 8 seconds longer, while the SVMMRF methods are about 100 seconds longer than the *MSSC-MSF* method. However, when presenting the processing time in Table 2.11, we have not taken into account the fact that after execution of the RHSEG segmentation algorithm, a segmentation map at the relevant level of hierarchy has been interactively chosen. Although this is a fast procedure, it is highly desirable to develop approaches for automatic selection of an appropriate hierarchical level.

Finally, since the incorporation of the spatial information in a classifier significantly improves accuracies, the pixelwise classification step can be speeded up by sacrificing some percentage of the classification accuracy. That can be achieved either by decreasing the number of pixels in the training set or by using another less time-consuming classifier than the SVM pixelwise classification.

As a conclusion, the proposed spectral-spatial classification approaches are also efficient in terms of computational complexity. It is of great interest to further consider possibilities of high-performance implementation of the developed techniques.

2.5.2 Parallel GPU processing of hyperspectral imagery

In the previous section, we have assumed that independent classification and segmentation tasks are executed at the same time on different processing units, thus speeding up the whole classification procedure. Hence, we have exploited *task-level parallelism*, referring to different and independent sets of instructions executing in parallel. Since many modern computers have several processors or processing elements, it is of interest to execute several tasks simultaneously.

Furthermore, as discussed in [112, 122] and in Paper 7, most of the hyperspectral image analysis algorithms exhibit inherent parallelism at multiple levels, and thus can be nicely mapped onto massively parallel processing systems. This significantly speeds up the processing time, which is an important issue for many applications, such as military defense, tracking and management of natural disasters (*e.g.*, forest fires), oil spills, and other types of chemical contamination, where timely classification is highly desirable. We have investigated the parallel implementation of a hyperspectral anomaly detection algorithm (*i.e.*, an unsupervised binary classification algorithm) using commodity Central Processing Units (CPUs) and GPUs. Paper 7 details the results of our study, including general principles of parallel high-performance computing on commodity processors and the structure of the developed algorithm implementations. Here, we briefly summarize approaches of parallelism and advantages of the GPU-based parallel processing of hyperspectral imagery.

2.5.2.1 Parallelizing algorithms

Plaza *et al.* [112] and Setoain *et al.* [122] have reviewed parallel processing of hyperspectral images. There are two main approaches to decompose the problem into parts that can be run concurrently: *task-level* decomposition and *data-level* decomposition [92, 102]. Setoain *et al.* [122] discriminate among task-level, spatial-level and spectral-level parallelism for the hyperspectral image processing algorithms (the last two levels are the particular cases of the data decomposition patterns).

Task-level parallelism is the characteristic of a parallel program to perform entirely different calculations on either the same or different sets of data. *Spatial-level* parallelism decomposes the image into subsets of pixel vectors that are operated on independently, thus forming data streams processed concurrently by the processing elements (the finest level being pixel-level decomposition, when each processing element is working on one pixel vector). *Spectral-level* parallelism refers to decomposition of a multi-band image data into units containing subsets of contiguous spectral bands.

As mentioned before, many hyperspectral data processing algorithms can be easily parallelized. Here, we give two examples of exploiting spatial-level and spectral-level parallelism of the methods proposed in the frame of this thesis:

Example 1 (*Pixelwise classification*): A pixelwise classifier assigns each pixel to one of the classes based on its spectrum. Consequently, the pixelwise classification task exhibits inherent parallelism at pixel level: After the classifier is trained, each pixel vector can be processed and classified independently from other pixels. The number of threads³ that can be run concurrently is equal to the number of pixels n . As the value of n is high, the amount of concurrency is significant.

Example 2 (*Multidimensional gradient*): As discussed in Section 2.2.1, a one-band gradient from the B -band image can be computed by considering this image as a set of B one-band images, and combining B gradients corresponding to B spectral channels into one image (for instance, using a sum operator). This approach is composed of two consequent steps, which can be both parallelized. 1) The first step consists in computing a gradient of every spectral channel (in total, B gradients). Hence, this task exhibits spectral-level parallelism: B threads can be executed concurrently, each of which computes

³A *thread* is a portion of a program that can run independently of and concurrently with other portions of the program.

a gradient for one band. Moreover, the task of computing a one-band gradient for one channel can be further parallelized (exploiting spatial-level decomposition). 2) In the second step, the value of each pixel in the resulting gradient image is computed by combining (for instance, summing) the values from B gradients at this pixel. Thus, n threads can be executed concurrently for computing the values of n pixels in the final gradient image (pixel-level parallelism).

Paper 7 gives more examples of decomposing hyperspectral data processing tasks into parts that can be executed in parallel. As a conclusion, hyperspectral image analysis algorithms possess a significant amount of data-level concurrency, suitable for a “single instruction multiple data” architecture that allows massively parallel processing.

2.5.2.2 High-performance parallel computing using GPUs

The parallel algorithms can be implemented and run onto supercomputers (such as Cray SV, Nec SX series) and cluster systems [122]. Unfortunately, there are two pitfalls in using these systems: 1) they are generally expensive; 2) they are too heavy and power-consuming for being embedded into onboard remote sensing systems.

An interesting recent development has been the introduction of fully programmable GPUs together with software interfaces like NVIDIA Compute Unified Device Architecture (CUDA) [105] and AMD CTM [6] dedicated to general purpose processing on video cards. Because the GPU architectures are optimized for massively parallel processing, modern commodity video cards can achieve very high computational performances for parallel problems, peaking at several hundred GFLOPS or more. The high demand for realistic graphics (and physics) in the computer game market drives the development of increasingly powerful GPUs at a low cost, while keeping computer architectures adapted to this technology to achieve very high bandwidth communication between the computer and the graphics hardware. Today, low-cost, low-weight gaming computers are readily available with extremely powerful parallel computing performance. This kind of hardware is therefore very well suited for on-board processing in a hyperspectral image analysis scenario.

Although General-Purpose computing on GPUs has been an active area of research for decades, the introduction of CUDA and CTM has finally brought it within reach of a broader community, giving programmers access to dedicated Application Programming Interfaces (APIs), Software Development Kits (SDKs) and GPU-enabled C programming language variants.

As a conclusion, four key advantages of parallel GPU processing of hyperspectral data can be distinguished:

1. *Speed*: Computational performances increase at a high rate through the generations of recent GPUs.
2. *Portability*: Compactness and suitable power consumption of GPUs makes possible their integration into onboard data processing systems.
3. *Low prices*: GPUs of the latest generation cost about US \$400-500 and become much cheaper as soon as the next generation appears.
4. *Programmability*: New hardware and software architectures, powerful APIs and SDKs are developed, facilitating the use of GPUs for parallel general-purpose computing.

We have used the CUDA technology to implement an anomaly detection algorithm on an Nvidia GeForce 8800 GPU, and compared its performance to a CPU-based implementation running on a dual quad-core computer (see Paper 7). The GPU provides a significant speedup of the algorithm, as compared to the CPU implementation. For some parts of the algorithm, speedups on the order of 10 and even 100 are observed. Thus, we have shown that the GPU enables real-time execution of the algorithm on a hyperspectral data stream with high spatial and spectral resolution, with acceptable detection performance and a significant margin on computing time. Therefore, it is of great interest to further explore parallel strategies for hyperspectral image analysis and classification, using commodity processors such as GPUs.

Chapter 3

Conclusions

THE main objective of this thesis was to propose and develop new methods and algorithms for hyperspectral data classification using spectral-spatial approaches. The focus was to incorporate spatial contextual information into the classification procedure for achieving high classification accuracies, while avoiding high computational loads. The next section summarizes the main contributions of this thesis, thus explaining how we met the objectives defined in Section 1.3. Then, perspectives for future work are discussed.

3.1 Contributions

Three general strategies for spectral-spatial classification of hyperspectral data have been proposed and investigated, all being efficient both in terms of classification accuracies and computational complexity. These three strategies have been presented in Sections 2.1 - 2.3 and in Papers 1-6 and can be summarized as follows:

1) The first strategy (described in Section 2.1) is based on the integration of the SVM method within an MRF framework for context classification. SVM and MRF models are two powerful tools for high-dimensional data classification and for contextual image analysis, respectively. Our proposed method consists in performing a probabilistic SVM pixelwise classification, followed by MRF-based regularization for incorporating spatial and edge information into the classification. An important novelty consists in defining and integrating the “fuzzy no-edge/edge” function into the spatial energy term involved in MRFs, which yields classification maps with accurate borders.

The proposed method has proven to be efficient and sufficiently robust for classifying different data sets. Although MRF regularization approaches are known to be time-consuming, appropriate optimization algorithms and modern computers enable one to execute the proposed approach within a short time interval. Hence, we recommend this technique for hyperspectral image classification, especially for images which do not contain large regions with uncertain spectral properties. If such regions exist, they risk to be misclassified by the pixelwise classifier, and the MRF-based method cannot reconstitute their true label. In this case, the segmentation-based context classification techniques are preferable.

2) In a second step (discussed in Section 2.2), we have proposed classification methods using adaptive spatial neighborhoods derived from region segmentation results. Different segmentation techniques (watershed, partitional clustering, and the RHSEG methods) have been investigated and extended to the case of hyperspectral images. We have concluded that the spatial-based segmentation techniques are preferable when an image contains classes with similar spectral responses, while spectral-based techniques are helpful if an image contains a lot of small structures and classes with dissimilar spectra. The segmentation techniques working both in the spatial and the spectral domain (for instance, the RHSEG method) are promising for obtaining accurate segmentation results on different images (though they are typically more time-consuming than the techniques working in either spatial or spectral domain only).

Then, approaches for combining the extracted spatial regions with spectral information in a classifier have been developed. In particular, a new spectral-spatial classification scheme is proposed, based on a pixelwise SVM classification, followed by majority voting within segmentation regions. This method has shown good classification results on different data sets and is the fastest among all the proposed spectral-spatial methods. We noticed that the choice of a relevant segmentation technique is important for achieving good classification performances.

3) In a third step (described in Section 2.3), we have concentrated on approaches to reduce oversegmentation in an image, which is achieved by automatically “marking” the spatial structures of interest before performing a marker-controlled segmentation. Our third proposed strategy is composed of two steps: The first step consists in analyzing probabilistic classification results for selecting the most reliably classified pixels as markers of spatial regions. Several marker selection methods have been proposed, using either individual classifiers, or a multiple classifier system. The second step is marker-controlled region growing. Different approaches have been developed for this purpose, using either watershed or Minimum Spanning Forest (MSF) methods and resulting in both segmentation and context classification maps.

Region growing segmentation based on automatically selected markers is a promising technique for context image analysis, since it mitigates dependence of segmentation performances on the chosen homogeneity criteria. The use of probabilistic classification results for marker selection has proven to be an efficient approach. We have concluded that it is advantageous to use an SVM classifier, spatial information and multiple classifier approaches in the marker selection procedure. The MSF-based method for marker-controlled region growing has proven to be more efficient and more robust technique when compared to the marker-controlled watershed. The proposed *MSSC-MSF* method (based on multiple spectral-spatial classifier marker selection followed by construction of an MSF techniques) has yielded the best or close to the best classification results for all the considered data sets and has proven to be computationally efficient. Thus, this method is highly recommended to be applied for context classification of hyperspectral data. Another recommended choice is the *SVMMSF+MV* classifier (using marker selection from probabilistic SVM results followed by construction of an MSF), which also yields high accuracies, is computationally efficient and easier to implement when compared to the *MSSC-MSF* method.

The developed classification methods have proven to efficiently handle high-dimensional data. Almost none of the developed techniques involves any feature extraction procedure (exceptions are the EM segmentation and the ML classification algorithms, which require dimensionality reduction).

Finally, we have investigated (in Section 2.5 and in Paper 7) possibilities of high-performance parallel computing using commodity CPUs and GPUs for hyperspectral image analysis and classification. In particular, we have used the CUDA technology to implement an anomaly detection algorithm on an Nvidia GPU, and compared its performances to a CPU-based implementation running on a dual quad-core computer. The GPU has proven to provide a significant speedup of the algorithm when compared to the CPU implementation. Section 3.3 gives a list of publications further explaining the contributions of this thesis.

3.2 Perspectives

The combination of segmentation and classification approaches for multi-band image analysis opens interesting perspectives. First of all, we have shown that the hierarchical segmentation technique yields accurate results. Unfortunately, the choice of a single (or several) segmentation level(s) is usually accomplished by means of supervised, or manual, procedures. Manual interaction is often subjective and time consuming. Therefore, it is of great interest to develop techniques for the automated selection of results in segmentation hierarchies.

Another important question concerns the choice of (dis)similarity measures in the image analysis procedures. Many spatial-based image processing techniques require to select a (dis)similarity measure between entities (pixels or regions), and performances of these techniques strongly depend on this choice.

In this thesis, we have explored the use of several measures (vector norms, SAM, SID measures) and we noticed that none of the tested measures have proven to be the most appropriate for different data sets or different techniques. This conclusion motivates further research on this topic and definition of novel (dis)similarity measures for hyperspectral data.

An important note is that in our second classification strategy, segmentation has been used for improving classification results. Our third classification strategy is based on performing probabilistic classification for selecting the most reliably classified pixels as markers for segmentation. Here, classification has helped for improving segmentation results. In all the cases, segmentation and classification techniques applied one after another, and their outputs are combined in some way. A further interesting research direction is to investigate how to combine segmentation and classification in the most efficient way for producing accurate maps of labeled objects. Thus, it would be of interest to interlace these procedures and develop a new segmentation-classification method which performs classification and segmentation concurrently, *i.e.*, applying a supervised classification technique during the construction of a segmentation map.

Advanced techniques for processing of hyperspectral data are often computationally expensive. We have shown the interest of high-performance parallel computing using commodity CPUs and GPUs for speeding up computational processes. It is of great interest to further explore parallel strategies for hyperspectral image analysis and classification, using commodity processors such as GPUs.

Finally, the proposed segmentation and classification techniques are applicable to various kinds of multivariate data. All of the strategies proposed in this thesis can be applied with similar excellent results to multispectral data. Moreover, it would be interesting to further extend and adapt the developed methods for particular applications, for instance: a) Urban mapping (information about shape structures and measurements of vegetation index can be included into classification system). b) Processing spectral-spatial information contained in hyperspectral data for target detection. c) Classification of medical images using spatial context.

3.3 List of Publications

Papers in Refereed International Journals (6):

1. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using Minimum Spanning Forest grown from automatically selected markers," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, DOI 10.1109/TSMCB.2009.2037132.
2. Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, 2010, DOI 10.1109/LGRS.2010.2047711.
3. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognition*, vol. 43, no. 7, pp. 2367-2379, July 2010.
4. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2973-2987, Aug. 2009.
5. Y. Tarabalka, T. V. Haavardsholm, I. Kåsen, and T. Skauli, "Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing," *Journal of Real-Time Image Processing*, vol. 4, no. 3, pp. 287-300, Aug. 2009.
6. Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Multiple spectral-spatial classification approach for hyperspectral data," *IEEE Trans. on Geoscience and Remote Sensing*, *under review*.

Book Chapter (1):

1. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral images using segmentation-derived adaptive neighborhoods," *Multivariate Image Processing* (edited by C. Collet, J. Chanussot, and K. Chendi), *ISTE Ltd and John Wiley & Sons Inc*, 2009, pp. 341-374.

International Conference Papers (8):

1. Y. Tarabalka, J. A. Benediktsson, J. Chanussot and J. C. Tilton, 2010. "A multiple classifier approach for spectral-spatial classification of hyperspectral data," in Proc. of *IGARSS'10*, Honolulu, Hawaii, USA, 2010.
2. J. Chanussot, J. A. Benediktsson, M. Fauvel, and Y. Tarabalka, "Spectral-spatial analysis in hyperspectral remote sensing: from morphological profiles to classified segmentation," in Proc of *SPIE*, Vol. 7477, 2009, pp. 1-14.
3. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Classification of hyperspectral images using automatic marker selection and Minimum Spanning Forest," in Proc. of *IEEE WHISPERS'09*, Grenoble, France, 2009, pp. 1-4.
4. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Classification based marker selection for watershed transform of hyperspectral images," in Proc. of *IGARSS'09*, Cape Town, South Africa, 2009, pp. III-105-III-108.
5. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Classification of hyperspectral data using Support Vector Machines and adaptive neighborhoods," in Proc. of *the 6th EARSeL SIG IS workshop*, Tel Aviv, Israel, 2009, pp. 1-6.
6. Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in Proc. of *IGARSS'08*, Boston, USA, 2008, pp. III-652-III-655.
7. Y. Tarabalka, T. V. Haavardsholm, I. Kåsen, and T. Skauli, "Parallel processing for normal mixture models of hyperspectral data using a graphics processor," in Proc. of *IGARSS'08*, Boston, USA, 2008, pp. II-990-II-993.
8. T. Skauli, I. Kåsen, T. Haavardsholm, A. Kavara, Y. Tarabalka, and Ø. Farsund, "Status of the Norwegian hyperspectral technology demonstrator," in Proc. of *NATO SET-130 "NATO Military sensing symposium"*, Orlando, Florida, USA, 2008, pp. F5-1-F5-6.

Résumé étendu

R.1 Introduction

LE concept d'imagerie hyperspectrale, ou spectroscopie, a été introduit dans les années 1980, quand A. F. H. Goetz et ses collègues du Jet Propulsion Laboratory de la NASA ont commencé une révolution en télédétection en développant de nouveaux instruments optiques comme le spectromètre imageur visible/infrarouge aéroporté AVIRIS [59, 111]. La *Spectrométrie* est définie comme “une acquisition simultanée d'images en de nombreuses bandes spectrales étroites et contiguës” [54]. Ainsi, l'imagerie hyperspectrale consiste en une mesure, un traitement et une analyse des spectres acquis d'une scène donnée à une distance courte, moyenne ou grande, par exemple avec un capteur aéroporté ou satellitaire. Les exemples de systèmes d'imagerie hyperspectrale sont AVIRIS [59], ROSIS (Spectromètre Imageur de Système Optique Réfléchissant) [50], HyMap [27], Hyperion [108]. Ils enregistrent typiquement des données dans une bande de 0.4 à 2.5 μm en utilisant de 115 à 512 canaux spectraux, avec une résolution spatiale de 0.75 à 30 m/pixel.

En imagerie hyperspectrale, chaque pixel peut être représenté comme un vecteur de haute dimension qui contient le spectre échantillonné de réflectance. Des matières différentes ayant des signatures spectrales différentes, l'imagerie hyperspectrale est une technologie bien adaptée pour de nombreuses applications de télédétection, y compris la classification précise des images. La *classification d'image hyperspectrale*, qui consiste en l'identification d'objets dans une scène acquise par un capteur imageur hyperspectral, est une tâche importante dans beaucoup de domaines appliqués tels que [82] :

- Agriculture de précision : La télédétection hyperspectrale est employée de manière intensive pour surveiller le développement et la santé des récoltes et ainsi empêcher le développement des maladies.
- Minéralogie : Beaucoup de minéraux peuvent être identifiés dans les images hyperspectrales, ce qui aide à rechercher des minéraux de valeur. Actuellement les chercheurs étudient l'effet des fuites de pétrole et de gaz des canalisations et des puits naturels sur les signatures spectrales de la végétation.
- Surveillance et gestion d'environnement : Les algorithmes de classification peuvent être employés pour créer les cartes thématiques afin d'analyser la croissance de zones urbaines et la sensibilité des différents secteurs de ville aux risques naturels, aussi bien que d'étudier des changements des écosystèmes terrestres ou en zones côtières et océaniques.
- Surveillance militaire : L'imagerie hyperspectrale est particulièrement utile dans la surveillance militaire, afin de détecter des objets qui sont invisibles pour l'oeil humain et ainsi fournir des services de sécurité.

Cependant, le grand nombre de canaux spectraux implique une grande dimensionnalité des données et complique ainsi l'analyse et la classification des images. Les premières méthodes d'analyse d'images hyperspectrales ont été basées sur des techniques existantes développées pour des images multispectrales, qui sont composées de quelques (d'habitude trois à sept) canaux spectraux. Cependant, la plupart des méthodes conçues et utilisées généralement pour l'analyse d'images en niveaux de gris, couleur ou multispectrale ne sont pas appropriées pour des images hyperspectrales. En effet, on a d'habitude une vérité de terrain ¹ avec un nombre très limité de pixels. Ceci limite les performances des techniques

¹Une *vérité de terrain* est définie comme des pixels étiquetés manuellement qu'on utilise pour l'apprentissage des classificateurs et pour l'évaluation de la précision de classification.

de classification supervisée. En outre, la malédiction de la dimensionnalité bien connue empêche des estimations statistiques robustes (ce problème est connu comme le phénomène de Hughes [66]). Par conséquent, il est nécessaire de développer des nouveaux algorithmes afin de profiter pleinement de toute la richesse de l'information spectrale.

Au cours de la dernière décennie, beaucoup de nouvelles méthodes pour la classification d'images hyperspectrales ont été proposées et étudiées. Ceci montre une croissance de l'intérêt pour l'analyse des données hyperspectrales. La plupart des méthodes de classification traitent chaque pixel indépendamment sans considérer les corrélations entre les pixels voisins (les méthodes de classification "pixel par pixel") [20, 53, 81]. En particulier, des Machines à Vecteurs de Support (Support Vector Machines : SVM) ont montré de bonnes performances en termes de précision de classification pour des données de grande dimension, y compris lorsque un nombre limité d'échantillons d'apprentissage est disponible [20, 43, 148].

Cependant, bien que la classification au niveau pixel permette la caractérisation précise des matériaux, elle ne prend pas en compte l'information contextuelle. Si on permute tous les pixels dans l'image d'une façon aléatoire, la classification de chaque pixel n'est pas changée, tandis que notre compréhension et interprétation visuelle sont sérieusement détériorées (on ne voit que le "bruit" après la permutation globale aléatoire). Cette expérience souligne l'importance de l'information contextuelle spatiale. Par conséquent, il est très important de développer des nouvelles techniques de classification qui sont capables de considérer les dépendances spatiales entre des pixels, afin d'améliorer les résultats de classification [43, 45, 74, 116, 138]. Une des méthodes de classification spectro-spatiale consiste à inclure l'information du voisinage afin de classifier chaque pixel. Les Champs Aléatoires de Markov (Markov Random Fields : MRFs) [42, 67], le fitrage morphologique [43, 45, 109] et les noyaux composés ont été étudiés à cette fin. Bien que les résultats de classification aient été améliorés par rapport à la classification pixel par pixel, l'utilisation de ces méthodes implique le problème du choix d'échelle ².

L'objectif général de cette thèse est de développer des nouvelles méthodes pour la classification des données hyperspectrales en utilisant conjointement l'information spectrale et spatiale. Dans la thèse, nous nous concentrons sur l'incorporation d'information contextuelle spatiale dans le processus de classification. Ceci soulève deux questions importantes : 1) Comment définir les structures spatiales, ou les voisinages pour chaque pixel, automatiquement ? 2) Comment combiner l'information spectrale et spatiale dans la classification ? Dans la suite, nous passerons en revue l'état de l'art pour la classification des images hyperspectrales. Ensuite, nous préciserons les objectifs de ce travail de thèse.

R.2 État de l'Art

R.2.1 Classification au niveau pixel

Le problème général de classification en imagerie hyperspectrale peut être décrit comme suit : On a un cube hyperspectral qui contient B bandes spectrales. Il peut être considéré comme un ensemble de n vecteurs $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$, chaque pixel étant représenté par un vecteur. Notons $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ un ensemble de classes d'intérêt dans la scène. La classification consiste à attribuer chaque pixel à une des K classes.

David Landgrebe et ses collègues dans la Laboratoire pour des Applications de la Télédétection à l'université de Purdue étaient les premiers à explorer les procédures pour l'analyse et la classification des données hyperspectrales. Ils ont adapté des techniques de reconnaissance des formes à cette fin [80, 81, 116]. Le schéma de classification qu'ils ont proposé est maintenant utilisé couramment pour la cartographie thématique des images hyperspectrales.

L'idée fondamentale de l'approche par reconnaissance des formes (pixel par pixel) est la suivante : Chaque pixel de l'image est considéré comme une forme (un modèle) et son spectre est considéré comme ensemble de caractéristiques. Puisque cet ensemble de caractéristiques est souvent redondant, il est fortement souhaitable d'effectuer une étape d'extraction/sélection de caractéristiques afin de réduire la

²Une *échelle* dans notre cas définit la taille des objets spatiaux d'intérêt.

dimension de l'ensemble des caractéristiques (de B dimensions dans les données de départ à D dimensions dans un nouvel espace de caractéristiques) et maximiser la séparabilité entre les classes.

Différentes techniques d'extraction de caractéristiques ont été proposées et étudiées à cette fin, telles que l'Analyse Discriminante (Discriminant Analysis Feature Extraction : DAFE), l'Extraction de Caractéristiques en utilisant une Frontière de Décision (Decision Boundary Feature Extraction : DBFE), et une méthode connue sous le nom de Non-parametric Weighted Feature Extraction (NWFE) [43, 81]. L'étape suivante ("*apprentissage*") consiste à diviser l'espace entier des caractéristiques en K régions de supports disjoints, de sorte que chaque point dans cet espace soit uniquement associé à l'une des K classes. Ainsi, chaque pixel est ensuite classifié selon son ensemble de caractéristiques.

Afin de diviser l'espace des caractéristiques en régions de décision, un ensemble d'échantillons représentatifs pour chaque classe, nommé comme "*des échantillons d'apprentissage*", est utilisé. Il est habituellement obtenu en marquant manuellement un nombre restreint de pixels dans une image. Ces échantillons d'apprentissage sont employés pour définir un modèle des classes dans l'espace des caractéristiques. Landgrebe a proposé d'utiliser des algorithmes de traitement des signaux à cette fin. Chaque classe ω_i peut être modélisée en termes de fonction de distribution de probabilité classe-conditionnelle (Probability Density Function : PDF) $p(\mathbf{x}|\omega_i)$. En supposant que chaque classe peut être décrite par une distribution (normale) Gaussienne [40], la classification par maximum de vraisemblance (Maximum Likelihood : ML) Gaussienne a été pendant beaucoup d'années le procédé standard de la cartographie thématique en télédétection hyperspectrale [116]. Essentiellement, il assigne chaque pixel à la classe ω_i en maximisant la probabilité a posteriori $p(\omega_i|\mathbf{x})$.

Cette méthode présente toutefois un désavantage important : si l'hypothèse sur la forme des PDFs classe-conditionnelles est erronée, les résultats de classification ne sont plus précis. En outre, le nombre élevé de caractéristiques disponibles, habituellement allié au nombre limité d'échantillons d'apprentissage, empêche l'évaluation fiable des paramètres statistiques. En conséquence, si un ensemble d'apprentissage est limité, la précision de classification tend à diminuer à mesure que la dimensionnalité augmente [80, 111].

Une classification ML plus simple consiste à modéliser chaque classe uniquement par sa réponse moyenne. Cette méthode, connue sous le nom de "*minimum distance to means classifier*" (algorithme de la distance minimale à la moyenne), assigne chaque pixel à la classe avec le plus proche vecteur moyen [80]. Une mesure de distance appropriée doit être choisie. L'algorithme du SAM (Spectral Angle Mapper), qui utilise la distance SAM afin de calculer la similitude des spectres dans l'espace multidimensionnel, est une des méthodes les plus simples de classification supervisée pour des données hyperspectrales [76]. La mesure SAM détermine la similitude spectrale entre deux vecteurs en calculant l'angle entre eux. Cette méthode ne fonctionne bien que dans le cas où la variabilité spectrale intraclasse est faible. Sinon, les classes ne peuvent pas être décrites par uniquement leurs vecteurs moyens, et la classification SAM échoue.

Dans les années 1990, les méthodes utilisant les réseaux de neurones ont suscité beaucoup d'attention dans le domaine de la classification d'images hyperspectrales [14, 95, 133, 154]. Les modèles des réseaux de neurones ont un avantage par rapport aux méthodes statistiques : ils n'ont pas besoin de la connaissance antérieure au sujet de la distribution statistique des classes. Un ensemble de poids et de non-linéarités décrivent le réseau de neurones. Ces poids sont calculés en utilisant un processus itératif d'apprentissage. L'intérêt pour ces méthodes a considérablement augmenté dans les années 1990, quand des nouvelles techniques d'apprentissage pour des données non-linéairement séparables ont été proposées [13]. Pourtant, l'utilisation des réseaux de neurones pour la classification d'images hyperspectrales a été limitée principalement à cause de leur complexité algorithmique. Vaiphasa [147] a proposé des algorithmes génétiques pour la classification des données hyperspectrales, qui sont capables de classer des échantillons non-linéairement séparables. Cependant, ces méthodes impliquent un calcul lourd.

Ces dernières années, on a proposé d'utiliser les SVM pour la classification d'images hyperspectrales et des noyaux afin de projeter les données dans un nouvel espace de caractéristiques où la séparabilité linéaire peut être exploitée. Les SVM se sont avérés extrêmement bien adaptés pour classer des données de grande dimension, quand un nombre limité des échantillons d'apprentissage est disponible [20, 43, 60, 148].

La méthode des SVM envisage de séparer des échantillons d'apprentissage appartenant à différentes classes en traçant les hyperplans de marge maximale dans l'espace où les échantillons sont projetés. Une analyse détaillée de la théorie des SVM et son utilisation pour la classification d'images hyperspectrales peut être trouvée dans [19, 60, 94, 148].

En conclusion, les SVM exploitent directement les propriétés géométriques des données, sans impliquer un procédé d'estimation de densité. Cette méthode s'est avérée beaucoup plus efficace que d'autres classifieurs non paramétriques (tels que les réseaux de neurones et un classifieur connu sous le nom de "k-Nearest Neighbor" (k-NN) [40]) en termes de précision de classification, complexité de calcul, et robustesse à la sélection des paramètres. Les SVM traitent d'une manière très efficace des données de grande dimension. Par conséquent, c'est une technique excellente pour éviter la procédure (habituellement assez complexe) de l'extraction/sélection de caractéristiques et ainsi simplifier le schéma traditionnel de classification proposé par Landgrebe [80]. En outre, les SVM montrent de bonnes possibilités de généralisation, exploitant entièrement les possibilités de discrimination à partir de relativement peu d'échantillons d'apprentissage. Tous ces avantages de la méthode des SVM en ont fait le classifieur le plus employé pour des données hyperspectrales dans la dernière décennie [86].

R.2.2 Classification spectro-spatiale

Les méthodes décrites jusqu'ici agissent au niveau pixel, puisqu'elles assignent chaque pixel à l'une des classes en n'analysant que ses propriétés spectrales, sans considérer l'information sur les pixels voisins. Cependant, dans une image réelle, les pixels voisins sont corrélés. Par conséquent, afin d'améliorer des résultats de classification, des méthodes de classification *spectro-spatiales*, ou des classifieurs basés sur le contexte, doivent être développés. Ils assignent chaque pixel de l'image à une classe en analysant : 1) ses propres valeurs spectrales (l'information spectrale) et 2) l'information extraite à partir de son voisinage (l'information spatiale). L'utilisation des classifieurs contextuels est particulièrement importante pour le traitement des images de haute résolution spatiale et/ou des données avec de grandes régions spatiales dans l'image. Dans la suite, nous décrivons les méthodes existantes pour la classification spectro-spatiale des données hyperspectrales.

Landgrebe et son équipe de recherche étaient des pionniers pour l'introduction du contexte spatial dans une classification d'images multibandes. Ils ont proposé le classifieur bien connu sous le nom de "ECHO" (Extraction and Classification of Homogeneous Objects) [74, 81], qui a été employé couramment en télédétection pour la classification spectro-spatiale des images multispectrales. Cette méthode est basée sur une croissance de régions afin de trouver des groupes homogènes des pixels voisins qui peuvent être caractérisés par des moyennes et des covariances de groupe. Chaque groupe est ensuite classifié par un classifieur ML.

Puisque la classification ECHO comporte une estimation des matrices de covariance, l'étape d'extraction de caractéristiques est conseillée avant d'appliquer cette technique aux données hyperspectrales (afin de réduire la dimension, et ainsi le nombre de paramètres à estimer). En conclusion, la méthode ECHO a introduit une approche puissante et largement utilisée pour la classification spectro-spatiale, qui consiste à segmenter une image en régions homogènes et ensuite classifier chaque région comme un objet entier.

La dernière décennie a été témoin des avancées importantes dans la classification contextuelle des images hyperspectrales. Un résumé de ces développements peut être trouvé dans [111]. La conception d'un système de classification spectro-spatiale implique deux étapes importantes :

1. La définition d'une approche pour *extraire l'information contextuelle spatiale* à partir de l'image. Par exemple, le classifieur ECHO segmente une image en des régions statistiquement homogènes à cette fin.
2. La définition d'une règle pour *combiner l'information spectrale et spatiale* dans un classifieur. Ici, on distingue deux catégories : (a) L'information spatiale est considérée dans l'étape de classification. Par exemple, l'information spectrale et spatiale peuvent être combinées dans un vecteur de caractéristiques pour chaque pixel, et ensuite la technique de classification pixel par pixel peut être appliquée à l'ensemble de ces vecteurs [22, 45]. Un autre groupe de méthodes de

cette catégorie distingue les objets dans l'image et ensuite classe chaque objet [74]. (b) Des dépendances spatiales sont considérées dans l'étape de décision. Un exemple est une classification pixel par pixel suivie de la régularisation spatiale de la carte de classification [16, 114, 117].

Ainsi, on doit généralement choisir une méthode (ou plusieurs méthodes) de classification des formes (par exemple, une méthode bayésienne, un réseau de neurones, une analyse discriminante linéaire (Linear Discriminant Analysis : LDA), SVM) pour l'employer dans le système de classification spectro-spatiale. Dans la suite, nous décrivons brièvement l'évolution des méthodes de classification spectro-spatiale pour des images hyperspectrales.

La manière la plus intuitive de définir le voisinage d'un pixel est de supposer que pour un pixel donné, ses voisins les plus proches appartiennent avec une probabilité élevée au même objet, *i.e.*, à la même classe [117]. Les voisinages composés de quatre ou huit pixels sont les plus souvent utilisés dans l'analyse d'image. Selon cette méthode, le pixel au centre peut être classifié en prenant en considération l'information de ses voisins.

En employant cette technique, Pony *et al.* [114] et Jackson et Landgrebe [67] ont proposé des classifieurs statistiques itératifs spectro-spatiaux pour des données hyperspectrales, qui utilisent une modélisation par MRFs. Les MRFs sont des modèles probabilistes qui sont appliqués couramment pour intégrer le contexte spatial dans des problèmes de classification d'image [39]. Dans la méthode proposée, la classification ML pixel par pixel est effectuée (la réduction de dimensionnalité est précédemment appliquée si nécessaire). Ensuite, la carte de classification est régularisée en utilisant le cadre de *maximum a-posteriori* (MAP)-MRF. Dans ce cadre, la règle de décision de MAP est formulée comme minimisation d'une fonction d'énergie qui comprend les termes d'énergie spectrale et spatiale [128]. Le terme d'énergie spectrale est dérivé de la probabilité classe-conditionnelle pour chaque pixel, alors que le terme d'énergie spatiale est calculé sur le voisinage de chaque pixel. En 2005, Farag *et al.* ont proposé une méthode similaire [42], qui emploie l'algorithme de régression de SVM basés sur les champs moyens. En conclusion, les méthodes utilisant MRFs se sont avérées être un outil puissant pour l'analyse contextuelle d'image. Cependant, ils comprennent une étape itérative d'optimisation, qui est longue en temps de calcul.

Dell'Acqua *et al.* ont appliqué le classifieur ARTMAP flou des réseaux de neurones pour la classification spectro-spatiale des images hyperspectrales [35, 51]. Les résultats de la classification pixel par pixel sont raffinés par une deuxième classification ARTMAP flou. Les vecteurs d'entrée pour la deuxième classification représentent les configurations de cartographie dans une fenêtre autour de chaque pixel. Cette méthode s'est avérée efficace dans des situations différentes. Cependant, elle a les mêmes désavantages que des autres techniques des réseaux de neurones (complexité de calcul importante et possibilités limitées de généralisation), qui sont devenues moins populaires ces dernières années.

Avec la popularité croissante des techniques des SVM pour l'analyse d'image hyperspectrale, il est intéressant d'incorporer le contexte spatial aux classifieurs SVM. Camps-Valls *et al.* ont proposé le cadre des noyaux composés à cette fin [22]. Dans un premier temps, pour chaque pixel, des vecteurs de caractéristiques spectrales et spatiales sont extraits séparément. Par exemple, la moyenne et l'écart type (par bande spectrale) peuvent être estimés sur le voisinage le plus proche du pixel et être employés comme des caractéristiques spatiales. Les vecteurs de caractéristiques spectrales et spatiales sont utilisés pour calculer deux matrices de noyau différentes. Ensuite, les deux matrices de noyau sont combinées en employant une des techniques proposées des noyaux composés, telle que la concaténation des vecteurs de caractéristiques, un noyau d'addition directe (direct summation kernel), un noyau d'addition pondérée (weighted summation kernel), un noyau "cross-information" et des noyaux pour la polyvalence améliorée. Cette méthodologie a montré de bonnes performances pour l'intégration de l'information contextuelle dans la classification et a été utilisée en considérant des systèmes de voisinage plus avancés.

En conclusion, l'avantage de l'approche des voisinages fixes est sa simplicité. On n'a pas besoin de calcul pour extraire les voisinages, puisqu'ils sont déjà définis par le système fixe. En outre, le voisinage inclut un nombre restreint de pixels, de ce fait simplifiant l'incorporation de l'information spatiale dans le classifieur. Les modèles probabilistes des MRFs sont employés couramment pour prendre en considération l'information de plus proche voisinage. De l'autre côté, les SVM ont été appliqués dans les techniques de classification spectro-spatiale les plus récentes. Par conséquent, il serait intéressant d'étudier à l'avenir

la combinaison des SVM et des techniques basées sur les MRFs pour la classification contextuelle.

Malheureusement, les voisinages fixes ne reflètent pas toujours l'information sur les structures spatiales d'une manière précise. Par exemple, ils provoquent l'assimilation des régions ne contenant que peu de pixels avec des structures voisines d'une taille plus importante et ils ne fournissent pas d'information spatiale précise à la frontière des régions.

Benediktsson *et al.* ont proposé d'utiliser les filtres morphologiques pour la classification spectro-spatiale [12]. La morphologie mathématique vise à analyser des relations spatiales entre les ensembles de pixels, *i.e.*, à extraire l'information sur la taille, la forme et l'orientation des structures. Ainsi, au lieu de définir un voisinage fixe pour chaque pixel, les filtres morphologiques permettent d'analyser le voisinage d'un pixel selon les structures auxquelles il appartient.

L'idée fondamentale de la morphologie mathématique est de comparer localement des structures dans l'image à une forme de référence, qu'on appelle l'Élément Structurant (Structuring Element : SE) [121, 126]. En utilisant des SEs de différentes formes et tailles, différentes structures dans l'image peuvent être analysées.

Afin d'analyser les structures spatiales dans des images de télédétection, Benediktsson *et al.* ont construit un **Profil Morphologique** (Morphological Profile : MP), qui se compose d'une image elle-même et deux ensembles d'images : un ensemble d'ouvertures et de fermetures par reconstruction, respectivement [12, 121, 126]. L'ouverture (resp. la fermeture) par reconstruction est un opérateur morphologique connexe satisfaisant l'affirmation suivante : Si une structure est plus claire (resp. plus foncée) que son entourage immédiat et ne peut pas contenir le SE, alors elle est totalement enlevée, sinon, elle est totalement préservée. Afin de déterminer les structures spatiales de différentes tailles, les SEs de différentes tailles sont employés pour calculer l'ensemble d'ouvertures (resp. fermetures) du MP.

On a d'abord proposé le MP pour des images en niveaux de gris [12]. Dell'Acqua *et al.* [35] ont appliqué la méthode de MP pour la classification des images hyperspectrales. À cette fin, la première Composante Principale (Principal Component : PC) des données multibandes est calculée [117] et l'image résultante est utilisée pour la construction d'un MP. Le MP est ensuite classifié par la méthode d'ARTMAP flou. Cependant, cette méthode n'a pas amélioré les résultats par rapport à la classification spectrale d'ARTMAP flou ou à la classification d'ARTMAP flou en utilisant les voisinages proches fixes.

En effet, en n'utilisant que la première PC pour construire le MP, on a réduit l'information spectrale d'une manière significative. Afin de traiter ce problème, Benediktsson *et al.* [11] ont proposé d'extraire plusieurs images qui contiennent des parties de l'information spectrale, et ensuite de construire un MP pour chacune des différentes images. Cette méthode s'appelle le **Profil Morphologique Étendu** (Extended Morphological Profile : EMP). Benediktsson *et al.* [11] ont construit des MPs pour les premières PCs des données hyperspectrales, et ont utilisé l'EMP comme un vecteur empilé pour la classification par réseaux de neurones. À la suite de ces travaux, Fauvel [43] a proposé d'exécuter l'extraction de caractéristiques pour les données hyperspectrales originales et pour l'EMP, et ensuite d'appliquer la classification SVM en utilisant les caractéristiques spectrales et spatiales concaténées [22]. Cette méthode permet de prendre davantage en compte l'information spectrale par rapport aux méthodes précédentes basées sur l'EMP.

Bien que les filtres d'ouverture et de fermeture morphologiques aient donné de bons résultats en extrayant les structures spatiales, ils ne peuvent pas fournir l'information spatiale complète pour une image, parce qu'ils ne traitent que les extrema de l'image. D'ailleurs, le même type d'objets peut être plus clair que leur voisinage dans certaines parties de l'image mais plus foncé dans d'autres. Afin d'analyser toutes les structures dans l'image, indépendamment de leurs valeurs de luminance (valeurs de gris), Soille a proposé d'utiliser les filtres auto-complémentaires [127]. Le **filtre de surface auto-complémentaire** enlève de petites structures de l'image en employant un critère de surface, et rapporte une carte des zones (composantes) connexes plates. Fauvel [43] a intégré ces filtres dans une nouvelle méthode de classification spectro-spatiale pour des images hyperspectrales. Chaque composante connexe de la carte obtenue par le filtrage de surface auto-complémentaire (le filtre doit être appliqué sur une image à une bande, par exemple sur la première PC, et la taille minimum d'une composante connexe doit être choisie par l'utilisateur) définit un voisinage adaptatif pour tous les pixels dans cette composante. Ensuite, le

vecteur médian [7] est calculé dans chaque voisinage, et est employé comme le vecteur de caractéristiques spatiales pour tous les pixels dans le voisinage considéré. Finalement, la classification SVM est effectuée en utilisant le noyau d'addition pondérée afin de combiner l'information spectrale et spatiale. Les résultats de classification sont comparables à ceux obtenus par la technique basée sur l'EMP [43].

En conclusion, les filtres morphologiques et de surface permettent une définition beaucoup plus précise du voisinage de chaque pixel et des structures spatiales dans l'image que les voisinages proches fixes. Ils sont particulièrement appropriés pour l'analyse des images de zones urbaines qui contiennent des structures (la plupart du temps d'origine humaine) avec une forme bien définie. Les exemples de ces objets sont les bâtiments, les routes et les arbres, qui peuvent être caractérisés par des formes rectangulaires, de ligne et de cercle, respectivement. Cependant, des opérateurs morphologiques sont à l'origine définis pour des images à une bande. Bien qu'on ait proposé leur extension au cadre hyperspectral, la plupart du temps en extrayant une ou quelques images mono-bandes représentatives et en traitant ces images indépendamment, la morphologie mathématique pour des images hyperspectrales doit être développée et définie de façon plus approfondie. On peut également noter que les méthodes décrites utilisent souvent des vecteurs de caractéristiques spectrales et spatiales afin de classifier chaque pixel. Ceci implique l'étape d'extraction de caractéristiques et une complexité de calcul importante. Ainsi, il est intéressant de développer de nouvelles techniques pour combiner l'information spatiale et spectrale dans un classifieur, qui surmonterait les inconvénients mentionnés. En outre, les voisinages spatiaux extraits par le filtrage morphologique ou de surface ne sont plus fixes, mais ils dépendent de *l'échelle choisie*. Les tailles de SEs pour le filtrage morphologique et le critère de surface pour le filtrage de surface doivent être choisis, et ils limitent les tailles des structures spatiales extraites. Ceci provoque des imprécisions dans l'information spatiale extraite.

Comment ces imprécisions peuvent-elles être davantage réduites ? Rappelons-nous comment le classifieur ECHO [74], une des premières techniques spectro-spatiales, extrait des voisinages spatiaux. L'algorithme ECHO divise une image en une partition de régions statistiquement homogènes, *i.e.*, effectue une **segmentation de l'image**.

La segmentation consiste à diviser l'image entière en des régions disjointes de sorte que chaque région soit considérée homogène selon un certain critère (critère d'homogénéité, par exemple, intensité ou texture) [69]. Ainsi, chaque région dans la carte de segmentation définit un voisinage spatial pour tous les pixels de cette région. Cette méthode est capable d'extraire de grands voisinages pour de grandes régions homogènes, tout en ne manquant pas des régions de la taille d'un ou quelques pixels. Ainsi, si une carte précise des régions, correspondant aux structures spatiales de l'image, pouvait être trouvée, la carte de segmentation correspondante fournirait des informations spatiales complètes et précises. Cependant, la segmentation automatique des images hyperspectrales est un problème difficile. La méthode ECHO est basée sur des calculs statistiques et comporte l'estimation des matrices de covariance. Par conséquent, cette technique n'est pas bien adaptée aux données hyperspectrales.

Aksoy [4] a proposé un autre schéma pour la classification spectro-spatiale de données hyperspectrales avec une étape de segmentation. Premièrement, des caractéristiques spectrales et spatiales sont extraites en employant LDA, l'Analyse en Composantes Principales (Principal Component Analysis : PCA) et le filtrage de Gabor [91]. Ces caractéristiques sont quantifiées et utilisées pour l'apprentissage des classifieurs Bayésiens avec les modèles de densité non paramétriques discrets. Ensuite, une segmentation est effectuée en utilisant l'algorithme itératif "split-and-merge". Cette technique applique un filtrage de surface sur la carte de classification obtenue au niveau pixel. Ensuite, les grandes régions de la carte de segmentation sont divisées en sous-régions plus petites et compactes. Finalement, les régions résultantes sont modélisées en utilisant leurs caractéristiques spectrales et spatiales (les bandes de Gabor et dix caractéristiques de forme pour chaque région, telles que la surface, l'excentricité, la compacité, et d'autres), et sont employées avec des classifieurs Bayésiens afin de calculer la carte de classification finale. Les excellents résultats de classification prouvent l'intérêt d'effectuer la segmentation d'image à partir des résultats de classification supervisée. Cependant, bien qu'Aksoy propose de définir les structures spatiales dans l'image en utilisant le cadre de la segmentation, la technique décrite de segmentation est basée sur le filtrage morphologique et est similaire à la méthode de Fauvel [43] qui emploie le filtrage de surface auto-complémentaire.

La méthode d'Aksoy enlève également les structures dont la surface est plus petite que le seuil choisi. En outre, cette méthode se compose de beaucoup d'étapes, et ainsi présente une complexité de calcul importante.

Linden *et al.* [88] ont appliqué la méthode des SVM pour la classification de chaque région dans une image hyperspectrale. Dans un premier temps, une segmentation par croissance de régions est effectuée, en utilisant le logiciel "eCognition" (les différents seuils mènent à différentes cartes de segmentation) [34]. Ensuite, un vecteur moyen pour chaque région est calculé. Finalement, les régions sont classifiées par un classifieur SVM. Les résultats expérimentaux obtenus ne produisent généralement pas une amélioration par rapport à ceux obtenus par la classification au niveau pixel. La raison peut se situer dans le fait que l'utilisation d'un vecteur moyen pour la classification de chaque région réduit de manière significative la quantité d'information spectrale présente dans la région. Par conséquent, il est intéressant d'explorer d'autres méthodes pour combiner l'information spatiale avec l'information spectrale afin d'améliorer les résultats de classification.

En conclusion, il est nécessaire de développer des nouvelles méthodes de classification contextuelle pour des images hyperspectrales, utilisant l'information spatiale extraite par segmentation. Les études récentes de Noyel [103] et Huang et Zhang [65], effectuées en même temps que ce travail de thèse, prouvent un grand intérêt pour ce sujet.

Noyel [103] a étudié la segmentation morphologique et la classification des images hyperspectrales. Tout d'abord, la dimensionnalité des données est réduite par une modélisation du spectre. Ensuite, la classification LDA pixel par pixel est effectuée. La carte de classification est filtrée classe par classe, en utilisant l'ouverture morphologique, afin de choisir de grandes régions spatiales comme des germes, ou des *marqueurs*, pour la segmentation par la méthode de la ligne de partage des eaux [56]. L'auteur a également proposé d'employer les boules aléatoires (les ensembles connexes de pixels des tailles aléatoirement choisies) extraites à partir de ces grandes régions comme des marqueurs. Puisque l'objectif final était de détecter les structures spécifiques, telles que des occlusions de colle et la croissance cancéreuse, des paramètres moyens spécifiques de chaque région ont été analysés afin d'obtenir la carte finale des objets extraits et du fond.

Huang et Zhang [65] ont récemment publié une étude comparative de plusieurs techniques spatiales pour la classification d'images hyperspectrales. Les auteurs ont proposé une nouvelle méthode basée sur une segmentation multi-échelle. Dans un premier temps, l'extraction de caractéristiques, connue sous le nom de Non-negative Matrix Factorization (NMF) est appliquée [83]. Ensuite, la segmentation d'image est effectuée, en utilisant la méthode "Mean Shift" (MS) ou la technique connue sous le nom de Fractal Net Evolution Approach (FNEA). La segmentation MS est une technique non paramétrique, qui détecte des modes d'une fonction de densité et divise une image dans des groupes de pixels en supposant que chaque mode soit un centre du groupe correspondant [29]. FNEA est une technique de fusion de régions, qui suppose que chaque pixel est une région, et ensuite elle fusionne avec procédure itérative des régions voisines satisfaisantes un critère d'homogénéité, jusqu'à la convergence [18]. Puisque les performances des deux techniques de segmentation dépendent des paramètres choisis (un critère d'homogénéité pour le FNEA et une largeur de bande spatiale pour la technique MS, respectivement), Huang et Zhang ont proposé d'utiliser les résultats de segmentation multi-échelle, obtenus en appliquant des algorithmes avec d différentes valeurs des paramètres. Ainsi, l'ensemble des caractéristiques spatiales pour un pixel donné contient d vecteurs moyens des régions qui contiennent le pixel considéré à différentes échelles de segmentation. Finalement, le classifieur des SVM est employé pour interpréter cet ensemble de caractéristiques. Cette technique multi-échelle a été comparée à la classification spectro-spatiale utilisant des SVM et des MPs dérivés (Derivative MPs : DMPs), la méthode étant proposée par Fauvel et testée pour des images panchromatiques [43]. Les excellents résultats de classification sont rapportés, où la technique basée sur les DMPs donne des résultats légèrement meilleurs que la méthode MS multi-échelle.

Ainsi, des techniques multi-échelles (Huang et Zhang [65]) et une segmentation par la méthode de la ligne de partage des eaux basée sur les marqueurs (Noyel [103]) sont conçues pour extraire l'information sur les structures spatiales de l'image, tout en évitant la dépendance forte des résultats à l'égard des paramètres de seuil. Bien que les méthodes multi-échelles fournissent les cartes de segmentation à dif-

férentes échelles et rapportent ainsi l'information spatiale détaillée, ils nécessitent une complexité de calcul importante et des grandes exigences de stockage. Tilton a proposé une technique de segmentation hiérarchique (Hierarchical SEGmentation : HSEG) pour des images hyperspectrales basée sur la croissance de régions et l'algorithme de groupement spectral. Cette technique donne également comme résultat un ensemble de segmentations d'image à différents niveaux de détails [143, 142]. Afin de réduire le temps de calcul, une approximation Récursive de HSEG (RHSEG) et sa version parallèle efficace ont été développées [142]. De l'autre côté, la segmentation par croissance de régions basée sur des marqueurs définis automatiquement est une technique prometteuse pour extraire les structures spatiales, puisqu'elle minimise la dépendance des résultats aux seuils choisis et le temps de calcul.

En conclusion, les méthodes de classification contextuelle ont prouvé leur avantage par rapport aux techniques de classification pixel par pixel, fournissant des cartes de classification plus précises et plus homogènes. Cependant, les techniques existantes souffrent des limitations d'échelle et/ou de complexité de calcul. Par conséquent, il est intéressant de développer des nouvelles méthodes et des nouveaux algorithmes efficaces et rapides pour l'analyse et la classification spectro-spatiale des images hyperspectrales.

R.3 Objectifs

L'**objectif principal** de cette thèse est le développement de nouvelles méthodes et de nouveaux algorithmes pour la classification spectro-spatiale des données hyperspectrales. Les approches développées devraient être efficaces en termes de précision de classification et de complexité de calcul. Après la discussion de la section précédente, nous avons défini les **objectifs** spécifiques suivants :

1. Combiner des approches de SVM et de MRFs pour la classification contextuelle des images hyperspectrales.
2. Développer des méthodes de classification utilisant l'information spatiale dérivée des résultats de segmentation. À cette fin, nous visons à
 - (a) Étudier des techniques de segmentation non-supervisée et les étendre au cadre hyperspectral, si nécessaire.
 - (b) Développer des règles pour combiner l'information spatiale extraite avec l'information spectrale afin de classifier une image.
 - (c) Concevoir des méthodes de classification utilisant la segmentation par croissance de régions basée sur des marqueurs définis automatiquement afin d'extraire les structures spatiales.
3. Essayer d'éviter le procédé habituellement long d'extraction/sélection de caractéristiques dans les nouvelles méthodes de classification. À cette fin, les techniques développées doivent pouvoir traiter des données de grande dimension. En particulier, le classifieur SVM est un bon choix de méthode de reconnaissance des formes.
4. Explorer des possibilités de calcul parallèle à haute performance sur les processeurs d'un usage commode (GPU) pour réduire une charge de calcul.

R.4 Contributions Principales

Les contributions principales de cette thèse sont récapitulées sur la Figure R.1, qui décrit la classification proposée des nouvelles méthodes pour la classification spectro-spatiale des données hyperspectrales. Afin d'atteindre les objectifs définis dans la section précédente, nous avons proposé et développé **trois stratégies générales** pour la classification des données hyperspectrales :

- 1) La première stratégie (décrite dans la Section 2.1 et dans l'Article 1) explore la classification spectro-spatiale utilisant des voisinages proches fixes et est basée sur l'intégration de la technique des SVM dans le cadre des MRFs. Notre méthode proposée consiste à effectuer une classification probabiliste des SVM, suivie de régularisation spatiale basée sur des MRFs. Une nouveauté importante consiste à intégrer la fonction de contour flou dans le terme d'énergie spatiale, afin de préserver des contours dans une carte de classification spectro-spatiale.

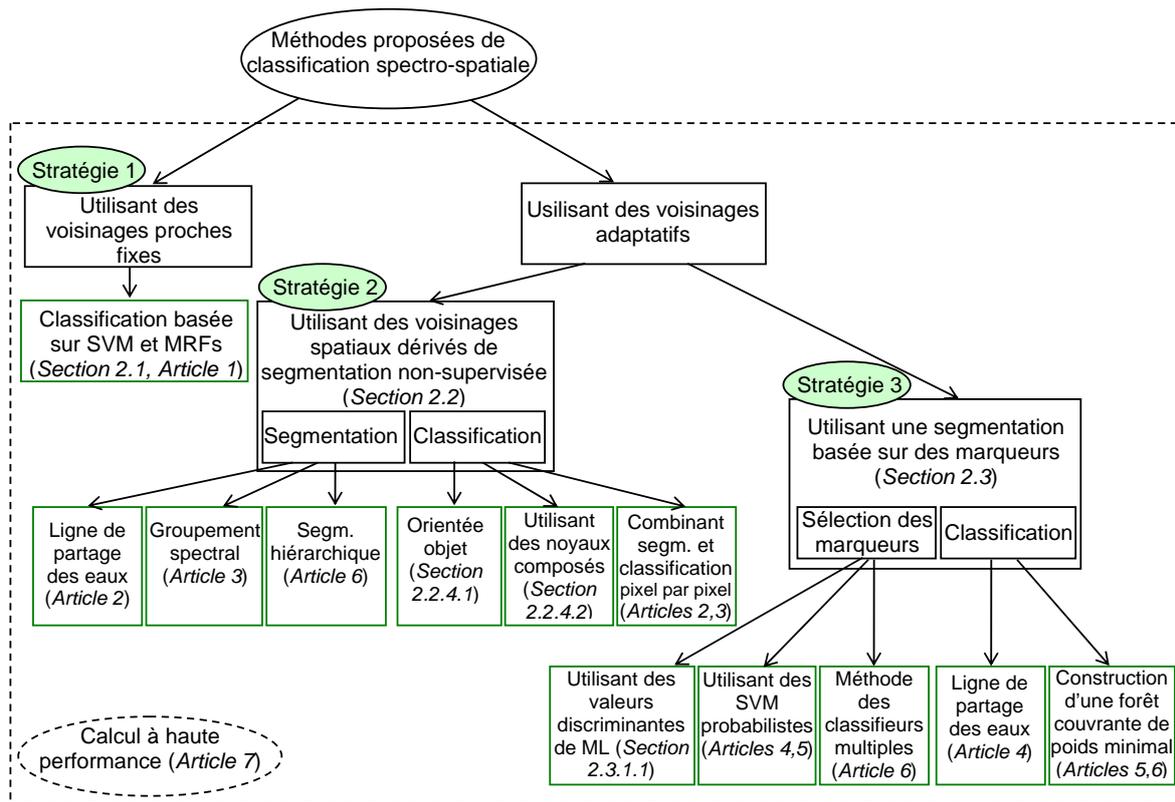


Figure R.1: Méthodes proposées de classification spectro-spatiale pour des données hyperspectrales (“segm.” signifie segmentation).

2) La deuxième stratégie (décrite dans la Section 2.2 et dans les Articles 2, 3, et 6) propose d’employer des voisinages spatiaux adaptatifs dérivés des résultats de segmentation. L’extension de différentes techniques de segmentation (les méthodes de la ligne de partage des eaux, du groupement spectral, et de RHSEG) au cadre hyperspectral est étudiée. Ensuite, des méthodes pour combiner les régions spatiales avec l’information spectrale dans un classifieur sont développées.

3) Dans la troisième stratégie (décrite dans la Section 2.3 et dans les Articles 4-6), nous avons proposé des techniques pour réduire la sur-segmentation. Ces méthodes utilisent des marqueurs des structures spatiales d’intérêt afin d’effectuer la segmentation par marqueurs. Une contribution importante consiste à analyser les résultats de la classification probabiliste afin de sélectionner les pixels les plus fiablement classés comme des marqueurs des régions spatiales. Nous avons proposé plusieurs méthodes pour la sélection des marqueurs, qui utilisent soit des classifieurs individuels, soit un ensemble de classifieurs. Ensuite, nous avons développé des techniques pour la segmentation par croissance de régions issues des marqueurs, en utilisant soit la ligne de partage des eaux, soit une forêt couvrante de poids minimal (minimum spanning forest), qui ont pour résultat les cartes de segmentation et de classification contextuelle.

Finalement, nous avons exploré (dans la Section 2.5 et dans l’Article 7) les possibilités du calcul parallèle à haute performance sur les processeurs d’un usage commode (GPU) afin de réduire la charge de calcul. Les méthodes développées se sont avérées efficaces en termes de précision de classification et de complexité de calcul, et ainsi montrent un grand potentiel pour les différents scénarios de l’analyse d’image.

R.5 Plan de la Thèse

Cette thèse se compose d'une collection des publications, avec le résumé étendu des recherches. Dans la première partie, le Chapitre 2 présente un résumé des méthodes et algorithmes développés dans le cadre de cette thèse pour la classification spectro-spatiale des images hyperspectrales. Le chapitre se termine par une comparaison expérimentale des méthodes proposées. Le Chapitre 3 conclut cette thèse en discutant les contributions principales et en suggérant des pistes de recherche. Finalement, cette partie se termine avec la bibliographie.

La Partie II se compose de sept publications qui présentent les résultats principaux de cette thèse :

Article 1. Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, 2010, DOI 10.1109/LGRS.2010.2047711.

Cet article présente une nouvelle méthode basée sur des SVM et des MRFs pour la classification spectro-spatiale des données hyperspectrales (Figure R.1, Stratégie 1). La méthode consiste à effectuer une classification probabiliste des SVM pixel par pixel, suivie de régularisation basée sur des MRFs afin d'incorporer l'information spatiale et de contour dans la classification.

Article 2. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognition*, vol. 43, no. 7, pp. 2367-2379, July 2010.

Dans cet article, nous proposons d'étendre l'algorithme de segmentation par ligne de partage des eaux pour des images hyperspectrales, afin d'extraire l'information sur les structures spatiales (Figure R.1, Stratégie 2). Les régions spatiales extraites sont ensuite utilisées en tant que voisinages adaptatifs pour la classification contextuelle. Le schéma de classification proposé est basé sur la classification des SVM pixel par pixel, suivie du vote majoritaire dans les régions de segmentation.

Article 3. Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2973-2987, Aug. 2009.

Cet article explore également l'utilisation des voisinages spatiaux adaptatifs dérivés des résultats de segmentation pour la classification des images hyperspectrales (Figure R.1, Stratégie 2). La méthode proposée combine les résultats d'une classification SVM pixel par pixel et une carte de segmentation obtenue par un algorithme de groupement spectral. Ceci est effectué en appliquant un vote majoritaire sur la classification spectrale pixel par pixel, en utilisant les voisinages adaptatifs définis par la carte de segmentation. L'utilisation de l'algorithme ISODATA et de la technique de résolution du modèle de mélange gaussien pour la segmentation d'image sont étudiées. Le bruit restant dans la carte de classification est ensuite réduit par un post-filtrage avec une fenêtre fixe.

Article 4. Y. Tarabalka, J. Chanussot, J. A. Benediktsson, "Classification based marker selection for watershed transform of hyperspectral images," in Proc. of *IGARSS'09*, Cape Town, South Africa, 2009, pp. III-105 - III-108.

Dans cet article, nous proposons d'analyser les résultats de la classification probabiliste afin de sélectionner les pixels classés de manière la plus fiable comme des marqueurs pour la segmentation par ligne de partage des eaux (Figure R.1, Stratégie 3). Chaque marqueur défini à partir des résultats de classification est associé à une étiquette de classe. Finalement, une carte de classification spectro-spatiale est obtenue en attribuant l'étiquette de classe de chaque marqueur à tous les pixels dans la région construite de ce marqueur.

Article 5. Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using Minimum Spanning Forest grown from automatically selected

markers,” *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, DOI 10.1109/TSMCB.2009.2037132.

Cet article présente une nouvelle méthode de classification, qui est basée sur la construction d’une forêt couvrante de poids minimal (Minimum Spanning Forest : MSF) à partir des marqueurs des régions (Figure R.1, Stratégie 3). Des marqueurs sont définis automatiquement à partir des résultats de classification probabiliste pixel par pixel. Chaque arbre dans la MSF forme une région dans la carte de segmentation. Une carte de classification est obtenue en attribuant une classe de chaque marqueur à tous les pixels dans la région construite de ce marqueur.

Article 6. Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, “Multiple spectral-spatial classification approach for hyperspectral data,” *IEEE Trans. on Geoscience and Remote Sensing*, *under review*.

Dans cet article, nous avons proposé une nouvelle méthode de classifieurs multiples pour la classification spectro-spatiale des images hyperspectrales (Figure R.1, Stratégie 3). Plusieurs classifieurs sont employés indépendamment pour classifier une image. Pour chaque pixel, si tous les classifieurs ont assigné ce pixel à la même classe, le pixel est considéré comme marqueur de la région spatiale, avec l’étiquette de la classe correspondante. Nous proposons d’utiliser des classifieurs spectro-spatiaux à l’étape préliminaire de la procédure de sélection des marqueurs, chacun d’eux combinant les résultats d’une classification pixel par pixel et une carte de segmentation. Les différentes méthodes de segmentation basées sur des principes différents mènent à différents résultats de classification. Finalement, une MSF est construite, où chaque arbre est enraciné sur un marqueur et forme une région dans la carte de classification.

Article 7. Y. Tarabalka, T. V. Haavardsholm, I. Kasen, and T. Skauli, “Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing,” *Journal of Real-Time Image Processing*, vol. 4, no. 3, pp. 287-300, Aug. 2009.

Cet article étudie l’utilisation des processeurs graphiques (Graphics Processing Units : GPUs) pour le traitement en temps réel de grands volumes de données hyperspectrales. En particulier, l’article étudie un algorithme de détection des anomalies basé sur la modélisation de mélange normal de la distribution spectrale de fond. La complexité et le potentiel pour la parallélisation de l’algorithme sont analysés, et les parties de l’algorithme qui possèdent une complexité de calcul importante sont implémentées sur un Nvidia GeForce 8800 GPU utilisant l’interface connu sous le nom de “Compute Unified Device Architecture programming interface”.

R.6 Évaluation expérimentale

Des résultats expérimentaux sont montrés dans la thèse sur quatre images aéroportées hyperspectrales, enregistrées par les capteurs AVIRIS et ROSIS, avec différents contextes (zones urbaines, agricoles et volcaniques), différentes résolutions spatiales (1.3 m et 20 m) et différents nombres de canaux spectraux (de 102 à 200 bandes). Cette section présente des résultats d’évaluation expérimentale sur une image de *Indian Pines*.

L’image *Indian Pines* est une zone de végétation, qui a été enregistrée par le capteur AVIRIS en Indiana du nord-ouest. L’image a des dimensions spatiales de 145×145 pixels, une résolution spatiale de 20 m par pixel, et une résolution spectrale de 200 bandes. Seize classes d’intérêt sont considérées, qui sont détaillées dans le Tableau R.1, avec le nombre d’échantillons pour chaque classe dans la vérité de terrain. L’image en fausses couleurs et la vérité de terrain sont présentées sur la Figure R.2. Nous avons aléatoirement choisi 50 échantillons pour chaque classe de la vérité de terrain comme échantillons d’apprentissage, sauf pour les classes “luzerne”, “herbe/pâturage-tondue” et “avoine”. Ces classes contiennent un faible nombre d’échantillons dans la vérité de terrain. Par conséquent, seulement 15 échantillons pour chacune de ces classes ont été aléatoirement choisis comme échantillons d’apprentissage. Les échantillons restants

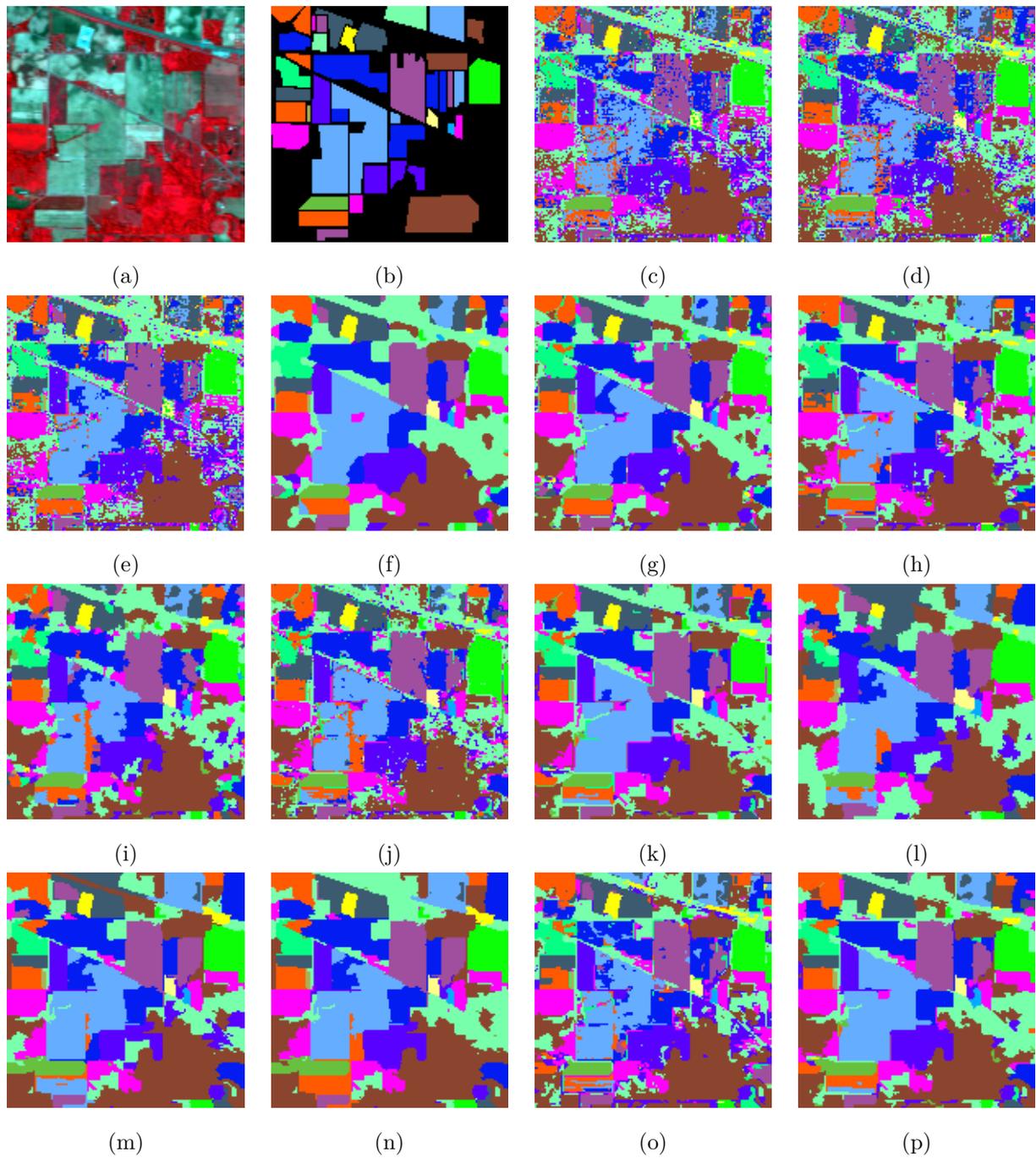


Figure R.2: Image de *Indian Pines*. (a) Image en fausses couleurs (bandes 837, 636, et 537 nm). Vérité de terrain : Maïs-no till, Maïs-min till, Maïs, Soja-no till, Soja-min till, Soja till, Luzerne, Herbe/pâturage, Herbe/arbres, Herbe/pâturage-tondue, Andain, Avoine, Blé, Bois, Btm-herbe-arbre, and Tours. (c-p) Cartes de classification. (c) ML. (d) SVM. (e) ECHO. (f) SVMMRF-NE. (g) SVMMRF-E. (h) MLMSF+MV. (i) WH+MV. (j) EM+MV. (k) RHSEG+MV. (l) SVM-WH. (m) SVMMSF. (n) SVMMSF+MV. (o) MC-MSF. (p) MSSC-MSF.

Table R.1: Précisions de classification pour l'image *Indian Pines* utilisant des méthodes pixel par pixel (approches spectrales) et spectro-spatiales, classes d'intérêt et nombre d'échantillons dans la vérité de terrain (No. d'Éch) : précision globale (Overall Accuracy : OA), précision moyenne (Average Accuracy : AA), coefficient Kappa (κ) et précisions pour chaque classe.

	No. d'Éch	Méthodes spectrales			ECHO	Méthodes de stratégie 1	
		3-NN	ML	SVM		SVMMRF-NE	SVMMRF-E
OA	-	66.27	75.41	78.17	82.64	92.05	91.83
AA	-	76.77	79.61	85.97	83.75	95.83	95.69
κ	-	62.04	72.25	75.33	80.38	90.93	90.71
1 - Maïs-no till	1434	41.84	71.39	78.18	83.45	93.28	98.48
2 - Maïs-min till	834	62.24	63.01	69.64	75.13	83.93	90.82
3 - Maïs	234	73.37	85.87	91.85	92.39	99.46	98.37
4 - Soja-no till	968	67.43	79.43	82.03	90.10	98.58	98.91
5 - Soja-min till	2468	53.91	52.65	58.95	64.14	82.09	76.92
6 - Soja till	614	64.72	85.99	87.94	89.89	97.70	97.34
7 - Luzerne	54	84.62	48.72	74.36	48.72	97.44	97.44
8 - Herbe/pâturage	497	86.35	93.51	92.17	94.18	97.54	97.54
9 - Herbe/arbres	747	91.97	94.69	91.68	96.27	97.70	97.56
10 - Herbe/pâturage-tondue	26	100	36.36	100	36.36	100	100
11 - Andain	489	95.67	97.72	97.72	97.72	99.54	99.54
12 - Avoine	20	80.00	100	100	100	100	100
13 - Blé	212	99.38	98.15	98.77	98.15	99.38	99.38
14 - Bois	1294	86.17	95.42	93.01	94.21	98.39	99.04
15 - Btm-herbe-arbre	380	45.15	73.03	61.52	81.52	88.18	79.70
16 - Tours	95	95.56	97.78	97.78	97.78	100	100

ont composé l'ensemble d'essai.

Nous comparons les méthodes suivantes pour la classification de l'image *Indian Pines* :

- Méthodes pixel par pixel (spectrales) : *3-NN*, *ML* et *SVM*.
- Méthodes spectro-spatiales proposées précédemment : *ECHO*.
- Méthodes de stratégie 1 (cf. Section 2.1) : *SVMMRF-NE* et *SVMMRF-E*. Classification probabiliste SVM, suivie d'une régularisation spatiale basée sur des MRFs, sans et avec la fonction de contour flou dans le terme d'énergie spatiale, respectivement.
- Méthodes de stratégie 2 (cf. Section 2.2 et Section 2.4) : *WH+MV*, *EM+MV* et *RHSEG+MV*. Segmentation par les méthodes de la ligne de partage des eaux, du groupement spectral, et de RHSEG, respectivement, suivie du vote majoritaire sur la classification spectrale pixel par pixel, en utilisant les voisinages adaptatifs définis par la carte de segmentation.
- Méthodes de stratégie 3 (cf. Section 2.3 et Section 2.4) :
 - *SVM-WH*. Sélection des marqueurs utilisant des SVM probabilistes, suivie de la construction de la ligne de partage des eaux.
 - *MLMSF* et *MLMSF+MV*. Sélection des marqueurs utilisant des valeurs discriminantes de ML suivie de la construction d'une forêt couvrante de poids minimal, sans et avec post-traitement, respectivement.

Table R.2: Précisions de classification pour l'image *Indian Pines* utilisant des méthodes basées sur une segmentation : précision globale (Overall Accuracy : OA), précision moyenne (Average Accuracy : AA), coefficient Kappa (κ) et précisions pour chaque classe.

	Méthodes de stratégie 2			Méthodes de stratégie 3						
	WH +MV	EM +MV	RHSEG +MV	SVM -WH	ML MSF	MLMSF +MV	SVM MSF	SVMMMSF +MV	MC- MSF	MSSC- MSF
OA	86.63	83.60	90.86	85.99	76.25	89.50	88.41	91.80	86.66	92.32
AA	91.61	85.34	93.96	86.95	82.32	91.45	91.57	94.28	92.58	94.22
κ	84.83	81.43	89.56	83.98	73.31	88.01	86.71	90.64	84.82	91.19
1	94.22	89.09	90.46	80.35	75.94	92.63	90.97	93.21	83.82	89.74
2	78.06	75.64	83.04	71.94	58.29	71.05	69.52	96.56	74.62	86.99
3	88.59	65.22	95.65	73.37	67.93	92.93	95.65	95.65	96.74	95.11
4	96.30	88.14	92.06	98.91	86.29	96.08	98.04	93.91	93.36	91.84
5	68.82	65.67	84.04	80.48	50.37	81.76	81.97	81.97	72.91	89.16
6	90.78	95.04	95.39	84.75	94.68	95.57	85.99	97.16	95.92	97.34
7	94.87	94.87	92.31	94.87	87.18	92.31	94.87	94.87	94.87	94.87
8	95.08	93.96	94.41	95.30	95.30	96.42	94.63	94.63	98.21	94.63
9	97.99	96.41	97.56	92.97	94.55	97.13	92.40	97.27	97.70	97.85
10	100	100	100	100	45.45	81.82	100	100	100	100
11	99.54	99.32	99.54	99.54	97.04	99.32	99.77	99.77	99.54	99.77
12	100	40.00	100	100	100	100	100	100	100	100
13	99.38	98.77	98.15	99.38	98.15	98.15	99.38	99.38	99.38	99.38
14	97.11	96.70	98.63	99.36	96.70	97.19	97.59	99.68	98.47	99.44
15	69.39	66.67	82.12	55.45	71.52	73.03	68.79	68.79	77.88	73.64
16	95.56	100	100	64.44	97.78	97.78	95.56	95.56	97.78	97.78

- *SVMMSF* et *SVMMSF+MV*. Sélection des marqueurs utilisant des SVM probabilistes, suivie de la construction d'une forêt couvrante de poids minimal, sans et avec post-traitement, respectivement.
- *MC-MSF* et *MSSC-MSF*. Sélection des marqueurs utilisant la méthode des classifieurs multiples (Multiple Classifiers : MCs) pixel par pixel et spectro-spatiaux, respectivement, suivie de la construction d'une forêt couvrante de poids minimal.

Le choix des paramètres pour ces méthodes est décrit dans la Section 2.4. Les Tableaux R.1 et 2.3 rapportent les précisions de classification globales : précision globale (Overall Accuracy : OA), précision moyenne (Average Accuracy : AA), coefficient Kappa (κ) et les précisions pour chaque classe pour les méthodes décrites. Les cartes de classification correspondantes sont montrées sur la Figure R.2. À partir de ces résultats et des résultats obtenus sur les autres images, les **conclusions** suivantes peuvent être tirées :

1. La méthode des SVM s'est avérée efficace pour la classification spectrale des données de grande dimension.
2. Il est avantageux de considérer des dépendances spatiales entre les pixels en effectuant la classification.
3. Les méthodes de classification spectro-spatiale développées dans le cadre de cette thèse donnent des meilleurs résultats que les méthodes ECHO et EMP proposées précédemment.
4. La régularisation basée sur les MRFs s'est avérée être un outil puissant pour l'analyse contextuelle des images. Les techniques de classification développées basées sur les MRFs (les méthodes *SVMMRF-NE* et *SVMMRF-E*) sont efficaces et suffisamment robustes pour classifier différents genres d'images. La technique *SVMMRF-E* inclut l'information de countour dans la régularisation contextuelle et fournit ainsi des cartes de classification avec des frontières précises.
5. En analysant plusieurs méthodes de segmentation non-supervisée, nous concluons que les techniques de segmentation dans le domaine spatial semblent être préférables lorsque l'image contient des classes avec des réponses spectrales similaires (par exemple, des images d'une zone de végétation, dans lesquelles on voudrait distinguer différentes espèces de végétation). Si une image contient des classes avec des réponses spectrales différentes, les techniques de segmentation dans le domaine spectral donnent de bons résultats. En outre, ces méthodes peuvent être utiles si une image contient beaucoup de structures petites et complexes qui risquent d'être assimilées avec de plus grandes régions voisines (par exemple, des images de zones urbaines). Finalement, les méthodes fonctionnant dans le domaine spatial et spectral sont prometteuses pour obtenir une segmentation précise de différentes images.
6. La méthode proposée de combinaison d'une carte de classification des SVM pixel par pixel et d'une carte de segmentation, utilisant une règle de vote majoritaire, s'est avérée être une technique efficace pour la classification contextuelle.
7. La segmentation par croissance de régions basée sur des marqueurs définis automatiquement est une technique prometteuse pour l'analyse contextuelle d'image. L'utilisation des résultats de classification probabiliste afin de sélectionner les pixels les plus fiablement classés comme des marqueurs des régions spatiales s'est avérée être une approche efficace. En particulier, il est avantageux d'utiliser un classifieur SVM, l'information spatiale et des techniques de MC dans la procédure de sélection de marqueurs. La méthode basée sur MSF pour une segmentation par marqueurs donne de bien meilleurs résultats que la technique de la ligne de partage des eaux contrôlée par marqueurs, et ainsi elle s'est avérée être une technique efficace et robuste de segmentation. La méthode *MSSC-MSF* a rapporté les meilleurs ou près des meilleurs résultats de classification pour toutes les images considérées et est ainsi fortement conseillée pour la classification contextuelle précise des données hyperspectrales.

R.7 Conclusions et Perspectives

L'objectif principal de cette thèse était de proposer et développer de nouvelles méthodes et de nouveaux algorithmes pour la classification des données hyperspectrales utilisant des approches spectro-spatiales.

Nous nous sommes concentrés sur l'incorporation d'information spatiale dans la procédure de classification afin d'obtenir des cartes de classification précises et d'éviter une charge de calcul importante. La section suivante résume les contributions principales de cette thèse, expliquant comment nous avons répondu aux objectifs définis dans la section R.3. Ensuite, nous discutons des perspectives et travaux futurs.

R.7.1 Contributions

Nous avons proposé et étudié trois stratégies générales pour la classification spectro-spatiale des données hyperspectrales, toutes étant efficaces en termes de précision de classification et de complexité de calcul. Ces trois stratégies sont présentées dans les Sections 2.1 à 2.3 et dans les Articles 1-6 et peuvent être récapitulées comme suit :

1) La première stratégie (décrite dans la Section 2.1) est basée sur l'intégration de la méthode des SVM dans le cadre des MRFs pour la classification contextuelle. Les SVM et les modèles des MRFs sont les deux outils efficaces pour la classification des données de grande dimension et pour l'analyse contextuelle d'images, respectivement. Notre méthode proposée consiste à effectuer une classification probabiliste SVM, suivie d'une régularisation basée sur les MRFs afin d'incorporer l'information spatiale et de contour dans la classification. Une nouveauté importante consiste à définir et intégrer la fonction de contour flou dans le terme d'énergie spatiale du modèle Markovien, qui permet d'obtenir des cartes de classification avec des contours précis.

La méthode proposée s'est avérée efficace et suffisamment robuste pour classifier différentes images. Bien que des approches de régularisation de MRF soient connues pour être longues en temps de calcul, les algorithmes d'optimisation appropriés et les ordinateurs modernes permettent d'exécuter la technique proposée sur un intervalle de temps court. Par conséquent, nous recommandons cette technique pour la classification des images hyperspectrales, particulièrement pour les images qui ne contiennent pas de grandes régions avec les propriétés spectrales incertaines. Si de telles régions existent, il y a un risque qu'elles ne soient pas classées correctement par le classifieur pixel par pixel. Dans ce cas-là, la méthode basée sur des MRFs ne peut pas retrouver leur classe correcte et les techniques de classification contextuelle basées sur une segmentation sont préférables.

2) Dans la deuxième stratégie (décrite dans la Section 2.2), nous avons proposé des méthodes de classification qui utilisent des voisinages spatiaux adaptatifs dérivés des résultats de segmentation. Les différentes techniques de segmentation (les méthodes de la ligne de partage des eaux, du groupement spectral, et de RHSEG) ont été étudiées et étendues au cadre hyperspectral. Nous avons conclu que les méthodes de segmentation dans le domaine spatial sont préférables si une image contient des classes avec des réponses spectrales similaires, alors que les techniques dans le domaine spectral sont utiles si une image contient beaucoup de petites structures et des classes avec des spectres différents. Les techniques fonctionnant dans le domaine spatial et spectral (par exemple, la méthode RHSEG) sont prometteuses pour la segmentation précise de différentes images (bien qu'elles soient en général plus longues en temps de calcul que les techniques ne fonctionnant que dans le domaine spatial ou spectral).

Ensuite, des approches pour combiner les régions spatiales extraites avec l'information spectrale dans un classifieur ont été développées. En particulier, nous avons proposé un nouveau schéma de classification, qui est basé sur la classification SVM, suivie du vote de majoritaire dans les régions de segmentation. Cette méthode a donné de bons résultats de classification sur différentes images et elle est la plus rapide parmi toutes les méthodes spectro-spatiales proposées. Nous avons noté que le choix d'une technique de segmentation appropriée est important afin d'obtenir de bonnes performances de classification.

3) Dans la troisième stratégie (décrite dans la Section 2.3), nous nous sommes concentrés sur des approches pour réduire la sur-segmentation dans une image, utilisant des marqueurs des structures spatiales définis automatiquement afin d'effectuer la segmentation par marqueurs. Notre troisième stratégie se compose de deux étapes : La première étape consiste à analyser les résultats de la classification probabiliste afin de sélectionner les pixels les plus fiablement classés comme des marqueurs des régions spatiales. Nous avons développé plusieurs méthodes pour la sélection des marqueurs, utilisant soit des classifieurs individuels, soit un système des classifieurs multiples. La deuxième étape est une croissance des régions

issues des marqueurs. Nous avons proposé différentes techniques à cette fin, qui utilisent soit la ligne de partage des eaux, soit une forêt couvrante de poids minimal (MSF) afin de construire des cartes de segmentation et classification contextuelle.

La segmentation par croissance des régions issues des marqueurs sélectionnés automatiquement est une technique prometteuse pour l'analyse d'image spectro-spatiale, puisqu'elle réduit la dépendance des performances de segmentation aux critères d'homogénéité. L'utilisation des résultats de classification probabiliste pour la sélection des marqueurs s'est avérée être une approche efficace. Nous avons conclu qu'il est avantageux d'utiliser un classifieur SVM, l'information spatiale et des méthodes des classifieurs multiples afin de sélectionner des marqueurs. La technique basée sur MSF pour la croissance de régions issues des marqueurs s'est avérée être une technique plus efficace et plus robuste que la méthode de la ligne de partage des eaux. La méthode *MSSC-MSF* (basée sur la sélection des marqueurs par la technique des classifieurs multiples spectro-spatiaux suivie de la construction de la MSF) a donné les meilleurs ou près des meilleurs résultats de classification pour toutes les images considérées et est efficace en termes de complexité de calcul. Ainsi, cette méthode est fortement conseillée pour la classification contextuelle des images hyperspectrales. Une autre technique conseillée est le classifieur *SVMMMSF+MV* (utilisant la sélection des marqueurs à partir des résultats des SVM probabilistes suivie de la construction de la MSF), qui est également efficace en termes de précision de classification et de complexité de calcul, et qui est plus facile à implémenter que la méthode *MSSC-MSF*.

Les méthodes de classification développées dans cette thèse se sont avérées traiter d'une manière efficace des données de grande dimension. Presque aucune des techniques proposées n'implique une étape d'extraction de caractéristiques (les exceptions étant la segmentation EM et la classification ML, qui nécessitent la réduction de dimension).

Finalement, nous avons étudié (dans la Section 2.5 et dans l'Article 7) les possibilités du calcul parallèle à haute performance utilisant les processeurs d'un usage commode (Central Processing Units : CPUs et GPUs) pour l'analyse et la classification des images hyperspectrales. En particulier, nous avons utilisé la technologie de CUDA afin d'implémenter un algorithme de détection des anomalies sur Nvidia GPU, et avons comparé ses performances avec l'implémentation sur CPU exécuté sur l'ordinateur de quadruple-noyau double (dual quad-core). L'implémentation sur GPU permet d'exécuter l'algorithme considéré beaucoup plus vite par rapport à l'implémentation sur CPU. La Section 3.3 donne une liste de publications qui constituent les contributions de cette thèse.

R.7.2 Perspectives

La combinaison des approches de segmentation et de classification pour l'analyse d'image multibande ouvre des perspectives intéressantes. Tout d'abord, nous avons montré que la technique de segmentation hiérarchique donne de bons résultats. Malheureusement, le choix d'un niveau de segmentation (ou plusieurs niveaux) est habituellement effectué au moyens de procédures supervisées, ou manuelles. L'interaction manuelle est souvent subjective et longue. Par conséquent, il est intéressant de développer des techniques pour le choix automatisé des résultats dans des hiérarchies de segmentation.

Une autre question importante concerne le choix des mesures de (dis)similarité dans les procédures d'analyse d'image. Beaucoup de techniques de traitement d'images fonctionnant dans le domaine spatial nécessitent d'utiliser une mesure de (dis)similarité entre pixels ou régions, et les performances de ces techniques dépendent fortement du choix de cette mesure. Dans cette thèse, nous avons exploré l'utilisation de plusieurs mesures (normes vectorielles, Spectral Angle Mapper : SAM, Spectral Information Divergence : SID) et nous avons noté qu'aucune des mesures considérées ne s'est avérée la plus appropriée pour différentes images ou différentes techniques. Ainsi, des nouvelles mesures de (dis)similarité pour des données hyperspectrales doivent être développées.

Notons que dans notre deuxième stratégie de classification, la segmentation a été utilisée afin d'améliorer les résultats de classification. Notre troisième stratégie est basée sur la sélection des marqueurs pour la segmentation en utilisant une classification probabiliste. Ici, la classification a aidé à améliorer les résultats de segmentation. Dans tous les cas, les techniques de segmentation et de classification se sont

appliquées l'un après l'autre, et leurs résultats sont combinées. Il serait intéressant d'entrelacer ces procédures et de développer une nouvelle méthode de segmentation-classification qui effectue la classification et la segmentation simultanément, *i.e.*, appliquant une technique de classification supervisée pendant la construction d'une carte de segmentation.

Les techniques avancées pour le traitement des données hyperspectrales possèdent souvent une charge de calcul importante. Nous avons montré l'intérêt du calcul parallèle à haute performance utilisant les CPUs et GPUs d'un usage commode afin de réduire le temps de calcul. Il est intéressant de continuer à explorer des stratégies parallèles pour l'analyse et la classification des images hyperspectrales, utilisant les processeurs d'un usage commode.

Finalement, les techniques de segmentation et de classification proposées peuvent être appliquées à divers genres de données multivariées. Par conséquent, il serait intéressant d'étendre et adapter les méthodes proposées pour des applications particulières, par exemple : a) Cartographie urbaine. b) Traitement d'information spectro-spatiale pour la détection des cibles dans les données hyperspectrales. c) Classification des images médicales utilisant le contexte spatial.

Bibliography

- [1] ITT Corporation, For further information about the ENVI software: <http://rsinc.com/envi/>.
- [2] N. Acito, G. Corsini, and M. Diani. An unsupervised algorithm for hyperspectral image segmentation based on the Gaussian mixture model. *Proc. of IGARSS'03*, 6:3745–3747, July 2003.
- [3] H. G. Akçay and S. Aksoy. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Trans. Geos. and Remote Sens.*, 46(7):2097–2111, July 2008.
- [4] S. Aksoy. Spatial techniques for image classification. In *C.H. Chen, ed., Signal and Image Processing for Remote Sensing, Taylor & Francis*, pages 491–513, 2006.
- [5] C. Allène, J. Audibert, M. Couprie, J. Cousty, and R. Keriven. Some links between min-cuts, optimal spanning forests and watersheds. In *Proc. of the 8th International Symposium on Mathematical Morphology*, pages 253–264, Rio de Janeiro, Brazil, October 2007.
- [6] AMD. ATI CTM guide, 2006.
- [7] J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. *Proc. of the IEEE*, 78(4):678–689, 1990.
- [8] G. Ball and D. Hall. ISODATA, a novel method of data analysis and classification. Technical report, Technical Report AD-699616, Stanford University, Stanford, CA, 1965.
- [9] J.-M. Beaulieu and M. Goldberg. Hierarchy in picture segmentation: a stepwise optimization approach. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(2):150–163, 1989.
- [10] J. A. Benediktsson and I. Kanellopoulos. Classification of multisource and hyperspectral data based on decision fusion. *IEEE Trans. Geos. and Remote Sens.*, 37(3):1367–1377, 1999.
- [11] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geos. and Remote Sens.*, 43(3):480–491, March 2005.
- [12] J. A. Benediktsson, M. Pesaresi, and K. Arnason. Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Trans. Geos. and Remote Sens.*, 41(9):1940–1949, Sept. 2003.
- [13] J. A. Benediktsson and P. H. Swain. *Statistical Methods and Neural Network Approaches for Classification of Data from Multiple Sources*. PhD thesis, Purdue Univ., School of Elect. Eng., West Lafayette, IN, 1990.
- [14] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy. Conjugate gradient neural networks in classification of very high dimensional remote sensing data. *Int. Journal Remote Sensing*, 14(15):2883 – 2903, 1993.

- [15] S. Beucher and C. Lantuejoul. Use of watersheds in contour detection. in *Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*. Rennes, France: CCETT/IRISA, Sept. 1979.
- [16] F. Bovolo and L. Bruzzone. A context-sensitive technique based on support vector machines for image classification. In *Proc. of PReMI*, pages 260–265, 2005.
- [17] G. Briem, J. A. Benediktsson, and J. R. Sveinsson. Multiple classifiers applied to multisource remote sensing data. *IEEE Trans. Geos. and Remote Sens.*, 40(10):2291–2299, Oct. 2002.
- [18] L. Bruzzone and L. Carlin. A multilevel context-based system for classification of very high spatial resolution images. *IEEE Trans. Geosc. and Remote Sens.*, 44(9):2587–2600, Sept. 2006.
- [19] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2:121–167, 1998.
- [20] G. Camps-Valls and L. Bruzzone. Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geos. and Remote Sens.*, 43(6):1351–1362, June 2005.
- [21] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon. Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection. *IEEE Trans. Geos. and Remote Sens.*, 46(6):1822–1835, June 2008.
- [22] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla. Composite kernels for hyperspectral image classification. *IEEE Geos. and Remote Sens. Letters*, 3(1):93–97, Jan. 2006.
- [23] B. C. Catanzaro, N. Sundaram, and K. Keutzer. Fast Support Vector Machine training and classification on graphics processors. Technical Report No. UCB/EECS-2008-11, 2008.
- [24] G. Celeux and G. Govaert. A Classification EM algorithm for clustering and two stochastic versions. *Comput. Stat. Data Anal.*, 14(3):315–332, 1992.
- [25] C. Chang and C. Lin. LIBSVM: A library for Support Vector Machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.
- [26] C. I. Chang. An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis. *IEEE Trans. Inf. Theory*, 46(5):1927–1932, 2000.
- [27] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields. The HyMap airborne hyperspectral sensor: The system, calibration and performance. In *Proc. of First EARSEL Workshop on Imaging Spectroscopy*, October 1998.
- [28] C. Collet, B. Pernet, and V. Mazet. Panoramic integral-field spectrograph: Ultraspectral data to understand the history of the universe. In *Multivariate Image Processing, Ch. Collet, J. Chanussot, and K. Chehdi (eds.), ISTE Ltd and John Wiley & Sons Inc*, pages 437–450, 2009.
- [29] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(5):603–619, May 2002.
- [30] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, Second Edition*. MIT Press and McGraw-Hill, 2001.
- [31] P. Corsini, B. Lazzerini, and F. Marcelloni. A fuzzy relational clustering algorithm based on a dissimilarity measure extracted from data. *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, 34(1):775–781, Feb. 2004.

- [32] C. Couprie, L. Grady, L. Najman, and H. Talbot. Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest. In *Proc. of ICCV*, pages 731–738, Kyoto, Japan, 2009.
- [33] N. Cristianini and J. Shawe-Taylor. *Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.
- [34] A. Darwish, K. Leukert, and W. Reinhardt. Image segmentation for the purpose of object-based classification. *Proc. of IGARSS '03*, 3:2039–2041, July 2003.
- [35] F. Dell'Acqua, P. Gamba, A. Ferrari, J.A. Palmason, J.A. Benediktsson, and K. Arnason. Exploiting spectral and spatial information in hyperspectral urban data with high resolution. *IEEE Geos. and Remote Sens. Letters*, 1(4):322–326, Oct. 2004.
- [36] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, B* 39(1):1–38, 1977.
- [37] E. Diday and J. C. Simon. Clustering analysis. In *Digital Pattern Recognition*, pages 47–94, 1976.
- [38] J. Driesen and P. Scheunders. A multicomponent image segmentation framework. In *Proc. of ACIVS*, pages 589–600, 2008.
- [39] R. C. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 20(5):121–154, 1993.
- [40] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley, 2001.
- [41] A.N. Evans and X.U. Liu. A morphological gradient approach to color edge detection. *IEEE Trans. Image Processing*, 15(6):1454–1463, June 2006.
- [42] A.A. Farag, R.M. Mohamed, and A. El-Baz. A unified framework for MAP estimation in remote sensing image segmentation. *IEEE Trans. Geos. and Remote Sens.*, 43(7):1617–1634, July 2005.
- [43] M. Fauvel. *Spectral and Spatial Methods for the Classification of Urban Remote Sensing Data*. PhD thesis, Grenoble Institute of Technology, 2007.
- [44] M. Fauvel, J. Chanussot, and J. A. Benediktsson. Adaptive pixel neighborhood definition for the classification of hyperspectral images with Support Vector Machines and composite kernel. In *Proc. of ICIP'08*, San Diego, USA, 2008.
- [45] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Trans. Geos. and Remote Sens.*, 46(11):3804–3814, Nov. 2008.
- [46] G. Flouzat, O. Amram, and S. Cherchali. Spatial and spectral segmentation of satellite remote sensing imagery using processing graphs by mathematical morphology. In *Proc. of IGARSS '98*, volume 4, pages 1769–1771, 1998.
- [47] G. M. Foody. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering & Remote Sensing*, 70(5):627–633, 2004.
- [48] K.S. Fu and J.K. Mui. A survey on image segmentation. *Pattern Recognition*, 13(1):3–16, 1981.
- [49] G. Fumera and F. Roli. A theoretical and experimental analysis of linear combiners for multiple classifier systems. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(6):942–956, June 2005.
- [50] P. Gamba. A collection of data for urban area characterization. In *Proc. of IGARSS'04*, volume I, pages 69–72, Anchorage, USA, Sept. 2004.

- [51] P. Gamba and F. Dell'Acqua. Increased accuracy multiband urban classification using a neuro-fuzzy classifier. *International Journal of Remote Sensing*, 24(4):827–834, 2003.
- [52] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, Nov. 1984.
- [53] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Comput. Electron. Agricult.*, 39:67–93, 2003.
- [54] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock. Imaging spectrometry for Earth remote sensing. *Science*, 228(4704):1147–1153, 1985.
- [55] O. Gómez, J. A. González, and E. F. Morales. Image segmentation using automatic seeded region growing and instance-based learning. In *Proc. of the 12th Iberoamericann Congress on Pattern Recognition*, pages 192–201, Valparaiso, Chile, November 2007.
- [56] R.C. Gonzalez and R.E. Woods. *Digital Image Processing, Second Edition*. Prentice Hall, 2002.
- [57] H. P. Graf, E. Cosatto, L. Bottou, I. Dourdanovic, and V. Vapnik. Parallel Support Vector Machines: The Cascade SVM. In *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2005.
- [58] 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. *IEEE Trans. Geos. and Remote Sens.*, 26(1):65–74, Jan 1988.
- [59] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. S. Solis, M. R. Olah, and O. Williams. Imaging spectroscopy and the Airborne Visible Infrared Imaging Spectrometer (AVIRIS). *Remote Sens. Environ.*, 65(3):227–248, September 1998.
- [60] J. A. Gualtieri and R. F. Cromp. Support Vector Machines for hyperspectral remote sensing classification. In *Proc. of SPIE*, volume 3584, pages 221–232, 1998.
- [61] M. R. Gupta and N. P. Jacobson. Wavelet Principal Component Analysis and its application to hyperspectral images. In *Proc. of IEEE ICIP'06*, pages 1585–1588, Oct. 2006.
- [62] T. Habib, J. Inglada, G. Mercier, and J. Chanussot. Speeding up Support Vector Machine (SVM) image classification by a kernel series expansion. In *Proc. of IEEE ICIP'08*, October 2008.
- [63] R. Haralick and L. Shapiro. Survey: Image segmentation techniques. *Computer Vision, Graphics, Image. Processing*, 29(1):100–132, 1985.
- [64] C. W. Hsu and C. J. Lin. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks*, 13:415–425, March 2002.
- [65] X. Huang and L. Zhang. A comparative study of spatial approaches for urban mapping using hyperspectral rosis images over pavia city, northern italy. *International Journal of Remote Sensing*, 30(12):3205–3221, 2009.
- [66] G. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Information Theory*, 14(1):55–63, Jan 1968.
- [67] Q. Jackson and D. Landgrebe. Adaptive bayesian contextual classification based on Markov random fields. *IEEE Trans. Geos. and Remote Sens.*, 40(11):2454–2463, November 2002.

- [68] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computing Surveys*, 31(3):264–323, 1999.
- [69] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*. McGraw-Hill series in Computer Science. McGraw-Hill, Inc., 1995.
- [70] Andrei C. Jalba, Michael H.F. Wilkinson, and Jos B. T. M. Roerdink. Automatic segmentation of diatom images for classification. *Microscopy Research and Technique*, 65(1-2):72–85, Sep 2004.
- [71] A.C. Jensen and A.S. Solberg. Fast hyperspectral feature reduction using piecewise constant function approximations. *IEEE Geos. and Remote Sens.*, 4(4):547–551, Oct. 2007.
- [72] L. O. Jimenez and D. A. Landgrebe. Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Trans. Systems, Man, and Cybernetics*, 28(1):39–54, February 1998.
- [73] L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data. An Introduction to Cluster Analysis*. John Wiley & Sons, 1990.
- [74] R. L. Kettig and D. A. Landgrebe. Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Trans. Geoscience Electronics*, 14(1):19–26, Jan. 1976.
- [75] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(3):226–239, Mar. 1998.
- [76] F. A. Kruse. Imaging spectrometer data analysis - a tutorial. In *International Symposium on Spectral Sensing Research (ISSSR)*, volume 1, pages 44–54, June 1994.
- [77] S. Kumar, J. Ghosh, and M. M. Crawford. Best-bases feature extraction algorithms for classification of hyperspectral data. *IEEE Trans. Geos. and Remote Sens.*, 39(7):1368–1379, Jul. 2001.
- [78] F. Lafarge, X. Descombes, and J. Zerubia. Textural kernel for SVM classification in remote sensing: application to forest fire detection and urban area extraction. In *Proc. of ICIP 2005*, volume 3, pages 1096–1099, Sept. 2005.
- [79] P. Lambert and L. Macaire. Filtering and segmentation: the specificity of color images. *Int. Conf. Color Graph. Image Process.*, 1:57–64, 2000.
- [80] D. Landgrebe. Hyperspectral image data analysis. *IEEE Signal Proces. Mag.*, 1053-5888:17–28, January 2002.
- [81] D. A. Landgrebe. *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley & Sons, Inc., 2003.
- [82] D. A. Landgrebe. Multispectral land sensing: Where from, where to? *IEEE Trans. Geos. and Remote Sens.*, 43(3):414–421, March 2005.
- [83] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 1999.
- [84] S. Lee and M. M. Crawford. Unsupervised multistage image classification using hierarchical clustering with a Bayesian similarity measure. *IEEE Trans. Image Processing*, 14(3):312–320, March 2005.
- [85] P. Li and X. Xiao. Evaluation of multiscale morphological segmentation of multispectral imagery for land cover classification. *Proc. of IGARSS'04*, 4:2676–2679, 20-24 Sept. 2004.

- [86] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba. Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRS-S data fusion contest. *IEEE Trans. Geos. and Remote Sens.*, 47(11):3857–3865, Nov. 2009.
- [87] H.-T. Lin, C.-J. Lin, and R. C. Weng. A note on Platt’s probabilistic outputs for Support vector Machines. Technical report, Department of Computer Science, National Taiwan University, 2003.
- [88] S. v. d. Linden, A. Janz, B. Waske, M. Eiden, and P. Hostert. Classifying segmented hyperspectral data from a heterogeneous urban environment using Support Vector Machines. *Journal of Applied Remote Sensing*, 1(1, 013543), 2007.
- [89] Desheng Liu, Maggi Kelly, and Peng Gong. A spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery. *Remote Sensing of Environment*, 101(2):167 – 180, 2006.
- [90] S. Makrogiannis, G. Economou, and S. Fotopoulos. A region dissimilarity relation that combines feature-space and spatial information for color image segmentation. *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, 35(1):44–53, Feb. 2005.
- [91] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(8):837–842, Aug. 1996.
- [92] T. G. Mattson, B. A. Sanders, and B. L. Massingill. *Patterns for Parallel Programming*. Addison Wesley, 2005.
- [93] F. Van Der Meer. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation*, 8(1):3–17, 2006.
- [94] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geos. and Remote Sens.*, 42(8):1778–1790, August 2004.
- [95] E. Merényi. Intelligent understanding of hyperspectral images through self-organizing neural maps. In *Proc. 2nd Int. Conf. Cybernetics and Information Technologies, Systems and Applications (CITSA 2005)*, pages 30–35, Orlando, FL, USA, 2005.
- [96] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [97] F. Meyer. Minimum spanning forests for image segmentation. In *Proc. Second International Conf. Math. Morphology and Its Applications to Image Processing*, pages 77–84, Sept. 1994.
- [98] F. Meyer. Topographic distance and watershed lines. *Signal Process.*, 38(1):113–125, 1994.
- [99] F. Meyer and S. Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1:21–46, 1990.
- [100] O. J. Morris, M. de J. Lee, and A. G. Constantinides. Graph theory for image analysis: an approach based on the shortest spanning tree. *IEE Proc. Communications, Radar and Signal Processing*, 133(2):146–152, April 1986.
- [101] S. Moussaoui, H. Hauksdottir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Douté, and J. A. Benediktsson. On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation. *Neurocomputing*, 71:2194–2208, 2008.
- [102] C. Nicolescu and P. Jonker. A data and task parallel image processing environment. *Parallel Computing*, 28(7-8):945–965, 2002.

- [103] G. Noyel. *Filtrage, réduction de dimension, classification et segmentation morphologique hyperspectrale*. PhD thesis, Center of Mathematical Morphology of the Paris School of Mines, 2008.
- [104] G. Noyel, J. Angulo, and D. Jeulin. Morphological segmentation of hyperspectral images. *Image Analysis & Stereology*, 26:101–109, 2007.
- [105] NVIDIA. NVIDIA CUDA compute unified device architecture - programming guide. <http://developer.nvidia.com/cuda>, June 2007.
- [106] V. Osma-Ruiz, J. I. Godino-Llorente, N. Sáenz-Lechón, and P. Gómez-Vilda. An improved watershed algorithm based on efficient computation of shortest paths. *Pattern Recognition*, 40(3):1078–1090, 2007.
- [107] P. Paclík, R.P.W. Duin, G.M.P. van Kempen, and R. Kohlus. Segmentation of multi-spectral images using the combined classifier approach. *Image and Vision Computing*, 21:473–482, 2003.
- [108] J. Pearlman, C. Segal, L. Liao, S. Carman, M. Folkman, B. Browne, L. Ong, and S. Ungar. Development and operations of the EO-1 Hyperion imaging spectrometer. In *Proc. of Earth Observing Systems V, SPIE*, volume 4135, pages 243–253, 2000.
- [109] M. Pesaresi and J. A. Benediktsson. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Trans. Geos. and Remote Sens.*, 39(2):309–320, Feb. 2001.
- [110] J. Platt. Probabilistic outputs for Support Vector Machines and comparison to regularized likelihood methods. In *A. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans, Advances in Large Margin Classifiers*, Cambridge, MA, 2000. MIT Press.
- [111] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113, Supplement 1:S110–S122, 2009.
- [112] A. Plaza, D. Valencia, J. Plaza, and P. Martínez. Commodity cluster-based parallel processing of hyperspectral imagery. *J. Parallel Distrib. Comput.*, 66(3):345–358, 2006.
- [113] A. J. Plaza and J. C. Tilton. Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images. In *Proc. of IGARSS '05*, volume 7, pages 4946–4949, July 2005.
- [114] O. Pony, X. Descombes, and J. Zerubia. Classification d'images satellitaires hyperspectrales en zone rurale et periurbane. Technical report, Sophia Antipolis, France, Sept. 2000.
- [115] R. C. Prim. Shortest connection networks and some generalizations. *Bell System Technology Journal*, 36:1389–1401, 1957.
- [116] J. A. Richards. Analysis of remotely sensed data: The formative decades and the future. *IEEE Trans. Geos. and Remote Sens.*, 43(3):422–432, March 2005.
- [117] J. A. Richards and X. Jia. *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag Berlin Heidelberg, 2006.
- [118] L. Rickard, W. Basedow, E. Zalewski, P. Silverglate, and M. Landers. HYDICE: An airborne system for hyperspectral imaging. In *Proc. of Imaging Spectrometry of the Terrestrial Environment, SPIE*, volume 1937, pages 173–179, 1993.
- [119] P. Scheunders. Multivalued image segmentation based on first fundamental form. In *Proc. of the 11th International Conference on Image Analysis and Processing*, pages 185–190, September 2001.

- [120] B. Scholkopf and A. J. Smola. *Learning with Kernels*. MIT Press, 2002.
- [121] J. Serra. *Image Analysis and Mathematical Morphology*. New York, Academic, 1982.
- [122] J. Setoain, C. Tenllado, M. Prieto, D. Valencia, A. Plaza, and J. Plaza. Parallel hyperspectral image processing on commodity graphics hardware. In *Proc. of ICPPW'06*, pages 465–472, Washington, DC, USA, 2006.
- [123] L. Shapiro and G. Stockman. *Computer Vision*. Prentice Hall, 2002.
- [124] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):888–905, Aug 2000.
- [125] P. Soille. Morphological partitioning of multispectral images. *Journal of Electronic Imaging*, 5(3):252–265, 1996.
- [126] P. Soille. *Morphological Image Analysis*. Springer-Verlag, 2 edition, 2003.
- [127] P. Soille. Beyond self-duality in morphological image analysis. *Image and Vision Computing*, 23(2):249–257, 2005.
- [128] A. H. S. Solberg, T. Taxt, and A. K. Jain. A Markov random field model for classification of multisource satellite imagery. *IEEE Trans. Geos. and Remote Sens.*, 34(1):100–113, Jan. 1996.
- [129] K. Staenz. Terrestrial imaging spectroscopy - some future perspectives. In *Proc. of the 6th EARSeL SIG IS workshop*, pages 1–12, Tel Aviv, Israel, 2009.
- [130] J. Stawiaski. *Mathematical Morphology and Graphs : Application to Interactive Medical Image Segmentation*. PhD thesis, Paris School of Mines, 2008.
- [131] B. Stevenson, R. O'Connor, W. Kendall, A. Stocker, W. Schaff, R. Holasek, D. Even, D. Alexa, J. Salvador, M. Eismann, R. Mack, P. Kee, S. Harris, B. Karch, and J. Kershenstein. The civil air patrol ARCHER hyperspectral sensor system. In *Proc. of Airborne ISR Systems and Applications II, SPIE*, volume 5787, pages 17–28, 2005.
- [132] S. Stoev. Rafsi - a fast watershed algorithm based on rainfalling simulation. In *Proc. of the 8th International Conference on Computer Graphics, Visualization, and Interactive Digital Media*, 2000.
- [133] S. Subramanian, N. Gat, M. Sheffield, J. Barhen, and N. Toomarian. Methodology for hyperspectral image classification using novel neural network. In *Proc. of SPIE.*, volume 3071, pages 128–137, Algorithms for Multispectral and Hyperspectral Imagery III, A. Evan Iverson; Sylvia S. Shen; Eds., August 1997.
- [134] S. Tadjudin and D. A. Landgrebe. Covariance estimation with limited training samples. *IEEE Trans. Geos. and Remote Sens.*, 37(4):2113–2118, July 1999.
- [135] W. Tao, H. Jin, and Y. Zhang. Color image segmentation based on mean shift and normalized cuts. *IEEE Trans. Systems, Man, and Cybernetics, Part B*, 37(5):1382–1389, Oct. 2007.
- [136] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot. Classification of hyperspectral data using Support Vector Machines and adaptive neighborhoods. In *Proc. of the 6th EARSeL SIG IS workshop*, pages 1–6, Tel Aviv, Israel, 2009.
- [137] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Trans. Geos. and Remote Sens.*, 47(9):2973–2987, Sept. 2009.

- [138] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers. *IEEE Trans. Systems, Man, and Cybernetics: Part B*, DOI: 10.1109/TSMCB.2009.2037132.
- [139] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Spectral-spatial classification of hyperspectral images using segmentation-derived adaptive neighborhoods. In *Multivariate Image Processing, Ch. Collet, J. Chanussot, and K. Chehdi (eds.), ISTE Ltd and John Wiley & Sons Inc*, pages 341–374, 2009.
- [140] 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.
- [141] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel. Segmentation and classification of hyperspectral data using watershed. In *Proc. of IGARSS'08*, pages III–652–III–655, Boston, USA, 2008.
- [142] J. C. Tilton. HSEG/RHSEG, HSEGViewer and HSEGReader user's manual (version 1.40). *Provided with the evaluation version of RHSEG available from: <http://ipp.gsfc.nasa.gov/RHSEG>*, 2008.
- [143] J.C. Tilton. Image segmentation by region growing and spectral clustering with a natural convergence criterion. In *Proc. of IGARSS '08*, volume 4, pages 1766–1768, 1998.
- [144] J.C. Tilton. Analysis of hierarchically related image segmentations. *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 60–69, 2003.
- [145] F. Tsai, C.-K. Chang, and G.-R. Liu. Texture analysis for three dimension remote sensing data by 3D GLCM. In *Proc. of 27th Asian Conference on Remote Sensing*, pages 1–6, August 2006.
- [146] Yu-Chang Tzeng. Remote sensing images classification/data fusion using distance weighted multiple classifiers systems. In *Proc. of PDCAT'06*, Taipei, Taiwan, Dec. 2006.
- [147] C. Vaiphasa. Innovative genetic algorithm for hyperspectral image classification. In *Proc. of the International Conference Map Asia*, volume 20, 2003.
- [148] V. Vapnik. *Statistical Learning Theory*. New York: Wiley, 1998.
- [149] L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Analysis and Machine Intel.*, 13(6):583–598, June 1991.
- [150] J. Wang and C.-I Chang. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Trans. Geos. and Remote Sens.*, 44(6):1586–1600, June 2006.
- [151] A. Widayati, B. Verbist, and A. Meijerink. Application of combined pixel-based and spatial-based approaches for improved mixed vegetation classification using IKONOS. In *Proc. of the 23th Asian conference on remote sensing*, 2002, 8 p.
- [152] T.-F. Wu, C.-J. Lin, and R. C. Weng. Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*, (5):975–1005, 2004.
- [153] L. Xu, A. Krzyzak, and C. Y. Suen. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. Systems, Man and Cybernetics*, 22(3):418–435, May/Jun 1992.
- [154] H. Yang, F. V. D. Meer, W. Bakker, and Z. J. Tan. A back-propagation neural network for mineralogical mapping from aviris data. *International Journal of Remote Sensing*, 20(1):97–110, 1999.

Part II

Publications

Paper 1

Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, “SVM and MRF-based method for accurate classification of hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, 2010, DOI 10.1109/LGRS.2010.2047711.

SVM- and MRF-Based Method for Accurate Classification of Hyperspectral Images

Yuliya Tarabalka, *Student Member, IEEE*, Mathieu Fauvel, Jocelyn Chanussot, *Senior Member, IEEE*, and Jón Atli Benediktsson, *Fellow, IEEE*

Abstract—The high number of spectral bands acquired by hyperspectral sensors increases the capability to distinguish physical materials and objects, presenting new challenges to image analysis and classification. This letter presents a novel method for accurate spectral-spatial classification of hyperspectral images. The proposed technique consists of two steps. In the first step, a probabilistic support vector machine pixelwise classification of the hyperspectral image is applied. In the second step, spatial contextual information is used for refining the classification results obtained in the first step. This is achieved by means of a Markov random field regularization. Experimental results are presented for three hyperspectral airborne images and compared with those obtained by recently proposed advanced spectral-spatial classification techniques. The proposed method improves classification accuracies when compared to other classification approaches.

Index Terms—Classification, hyperspectral images, Markov random field (MRF), support vector machine (SVM).

I. INTRODUCTION

HYPERSPECTRAL imaging sensors measure the energy of the received light in tens or hundreds of narrow spectral bands in each spatial position in the image [1]. Thus, every pixel can be represented as a high-dimensional vector across the wavelength dimension, called the spectrum of the material in this pixel. Since different substances exhibit different spectral signatures, hyperspectral imagery is a well-suited technology for accurate image classification, which is an important task in many application domains (monitoring and management of the environment, precision agriculture, etc.).

Most classification methods process each pixel independently without considering the correlations between spatially adjacent pixels (so-called *pixelwise* classifiers) [2], [3]. In

particular, support vector machines (SVMs) have shown good performances for classifying high-dimensional data when a limited number of training samples are available [3], [4]. Furthermore, spatial contextual information should help for an accurate scene interpretation. Therefore, it is very important to develop *spectral-spatial* classification techniques that are capable to consider spatial dependences between pixels [5]–[8].

In general, two categories of spectral-spatial classification methods can be distinguished. First, spatial contextual information is exploited in the classification stage. For instance, spectral and spatial information can be combined within a feature vector of each pixel, and then, a pixelwise classification technique can be applied to the obtained set of vectors [6], [9]. Another group of methods from this category first defines the objects within the image scene and then classifies each object [2], [5]. Second, spatial dependences are considered in the decision rule [10]. An example is a pixelwise classification followed by spatial regularization of the classification map.

Markov random fields (MRFs) are probabilistic models that are commonly used to integrate spatial context into image classification problems [7], [10], [11]. In the MRF framework, the *maximum a posteriori* (MAP) decision rule is typically formulated as the minimization of a suitable energy function [12]. An extensive literature is available on MRF-based image classification techniques. In particular, the research groups of Farag [7], Bruzzone [10], and Gong [11] have investigated the integration of the SVM technique within an MRF framework for accurate spectral-spatial classification of remote sensing images. All of them use SVMs to estimate class conditional probability density functions and MRFs to estimate context-based class priors. Farag *et al.* [7] have applied the mean field-based SVM regression algorithm for density estimation, with the purpose of hyperspectral image classification. Good classification results are reported, although no comparison with other advanced spectral-spatial classification techniques is published.

This letter presents a novel SVM- and MRF-based (SVMMRF) method for spectral-spatial classification of hyperspectral images. In the first step of the proposed method, a probabilistic SVM pixelwise classification of the hyperspectral image is applied. In the second step, spatial contextual information is used for refining the classification results obtained in the first step. This is achieved by means of the MRF regularization. An important difference from previously proposed methods [7], [10], [11] consists in defining and integrating the “fuzzy no-edge/edge” function into the spatial energy function involved in MRFs, aiming at preserving edges while performing spatial regularization.

Manuscript received January 25, 2010; revised March 16, 2010. Date of publication May 13, 2010. This work was supported in part by the Marie Curie Research Training Network “HYPER-I-NET.”

Y. Tarabalka is with the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA Lab), Grenoble Institute of Technology, 38402 Grenoble, France, and also with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (email: yuliya.tarabalka@hyperinet.eu).

M. Fauvel is with the Modelling and Inference of Complex and Structured Stochastic Systems (MISTIS) Team, National Institute for Research in Computer Science and Control (INRIA), 38334 Saint Ismier, France (e-mail: Mathieu.fauvel@inrialpes.fr).

J. Chanussot is with GIPSA Lab, Grenoble Institute of Technology, 38402 Grenoble, France (e-mail: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr).

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2010.2047711

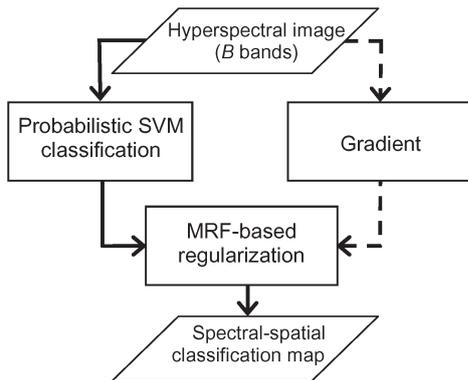


Fig. 1. Flowchart of the proposed SVMMRF classification scheme.

The second contribution of this letter consists in the experimental comparison of the presented approach with other recently proposed advanced spectral-spatial classification techniques. Experimental results are demonstrated on three hyperspectral airborne images recorded by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and the Reflective Optics System Imaging Spectrometer (ROSIS).

The outline of this letter is as follows. In the next section, an SVMMRF classification scheme for hyperspectral images is presented. Experimental results are discussed in Section III. Finally, conclusions are drawn in Section IV.

II. SVMMRF CLASSIFICATION SCHEME

The flowchart of the proposed SVMMRF classification method is shown in Fig. 1. At the input, a B-band hyperspectral image is given, which can be considered as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. Let $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$ be a set of information classes in the scene. Classification consists in assigning each pixel to one of the K classes of interest.

A. Probabilistic SVM Classification

The first step of the proposed procedure consists in performing a probabilistic SVM pixelwise classification of the hyperspectral image [4], [13]. Other probabilistic classifiers could be used. However, SVMs are extremely well suited to classify hyperspectral data [3]. The standard SVMs do not provide probability estimates for the individual classes. In order to get these estimates, pairwise coupling of binary probabilistic estimates is applied [13], [14].

B. Computation of the Gradient

Independent of the previous step, a one-band gradient of the hyperspectral image is computed, which is further used for defining the fuzzy no-edge/edge function. Approaches for defining a one-band gradient from the B-band image are analyzed in [15]. Here, we first compute horizontal, vertical, and two diagonal gradients (corresponding to the directions 0° , 90° , 45° , and 135° , respectively), using Sobel masks [16], where each of the gradients is computed as the sum of the gradients of every spectral channel. The resulting one-band gradient

$\nabla(\mathbf{X}) = \{\rho_j \in \mathbb{R}, j = 1, 2, \dots, n\}$ is found as the average of the four obtained directional gradients.

C. MRF-Based Regularization

In the final step, the regularization of the SVM classification map is performed, using the MAP-MRF framework. This framework is based on the interpixel class dependence assumption, which means that a pixel belonging to a class ω_i is likely to have neighboring pixels belonging to the same class. In our work, an eight-neighborhood is assumed (let \mathbf{N}_i be the set of neighbors for a given pixel \mathbf{x}_i).

We adopt the Metropolis algorithm, based on stochastic relaxation and annealing, for computing the MAP estimate of the true classification map given the initial (pixelwise) classification map [17], [18]. The considered method is based on the Bayesian approach and aims at minimizing the global energy in the image, by iterative minimization of local energies (defined hereafter) associated with randomly chosen image sites, i.e., pixels.

Let $\mathbf{L} = \{L_j, j = 1, 2, \dots, n\}$ be a generic set of information class labels for the image \mathbf{X} . We propose to compute the local energy of a given site associated with a pixel \mathbf{x}_i as

$$U(\mathbf{x}_i) = U_{\text{spectral}}(\mathbf{x}_i) + U_{\text{spatial}}(\mathbf{x}_i) \quad (1)$$

where $U_{\text{spectral}}(\mathbf{x}_i)$ is the spectral energy function from the observed data and $U_{\text{spatial}}(\mathbf{x}_i)$ is the spatial energy term computed over the local neighborhood \mathbf{N}_i . We define the spectral energy term as

$$U_{\text{spectral}}(\mathbf{x}_i) = -\ln\{P(\mathbf{x}_i|L_i)\} \quad (2)$$

where $P(\mathbf{x}_i|L_i)$ is estimated by pairwise coupling of probability estimates from “one-versus-one” SVM outputs [11], [14].

For the spatial energy term, two different expressions are investigated. We first consider the standard spatial energy expression, used, for instance, in [10], which is computed as

$$U_{\text{spatial}}^{\text{NE}}(\mathbf{x}_i) = \sum_{\mathbf{x}_j \in \mathbf{N}_i} \beta(1 - \delta(L_i, L_j)) \quad (3)$$

where $\delta(\cdot, \cdot)$ is the Kronecker delta function ($\delta(a, b) = 1$ if $a = b$, and $\delta(a, b) = 0$ otherwise) and β is a parameter that controls the importance of the spatial versus spectral energy terms. The superscript “NE” means that no edge information is taken into account. The term $U_{\text{spatial}}^{\text{NE}}(\mathbf{x}_i)$ is proportional to the number of neighboring pixels of \mathbf{x}_i assigned to one of the classes different from L_i . This spatial energy term is particularly suitable for the images with large spatial structures. However, if a small one-pixel object is present in the image, this model will favor assigning this pixel to the class of the surrounding objects.

In order to mitigate this drawback of the previous spatial term and to preserve small structures and edges in the classification map, we propose to integrate the edge information into the spatial energy function. The computation of an accurate edge map for hyperspectral images is a challenging task. For instance, it can be obtained by thresholding the gradient image $\{\rho_j \in \mathbb{R}, j = 1, 2, \dots, n\}$. For this purpose, an appropriate threshold

must be chosen. Instead of computing the edge map, we propose to define the following “fuzzy no-edge/edge function”:

$$\varepsilon(\mathbf{x}_j) = 1 - \frac{\rho_j}{\alpha + \rho_j} \quad (4)$$

where α is a parameter controlling the approximate edge threshold. From here, the following spatial energy function is proposed:

$$U_{\text{spatial}}^E(\mathbf{x}_i) = \sum_{\mathbf{x}_j \in \mathbf{N}_i} \beta \varepsilon(\mathbf{x}_j) (1 - \delta(L_i, L_j)). \quad (5)$$

The superscript “ E ” means that the edge information is taken into account. In the following, we thus refer to two different methods, namely, SVMRF-NE and SVMRF-E, when (3) and (5) are used for computing the spatial energy, respectively.

We briefly summarize the considered Metropolis algorithm for optimizing the energy function. In each iteration, an image site (i.e., a pixel \mathbf{x}_i) is randomly chosen. The local energy of the given site $U(\mathbf{x}_i)$ is computed by (1). Then, a new class label L_i^{new} is randomly selected for the site \mathbf{x}_i , and a new local energy $U^{\text{new}}(\mathbf{x}_i)$ is computed. If the variation of the energy $\Delta U = U^{\text{new}}(\mathbf{x}_i) - U(\mathbf{x}_i) < 0$, the new class label is assigned to \mathbf{x}_i : $L_i = L_i^{\text{new}}$. Otherwise, the new class assignment is accepted with the probability $p = \exp(-\Delta U/T)$. Here, T is a global control parameter called “temperature” [18]. The optimization begins at a high temperature, which is gradually lowered as the relaxation procedure proceeds. This procedure avoids converging to local *minima*.

III. EXPERIMENTAL RESULTS AND DISCUSSION

We applied the proposed SVMRF-NE and SVMRF-E classification approaches to three hyperspectral airborne images described in the following:

- 1) The *Indian Pines* image is of a vegetation area that was recorded by the AVIRIS sensor. The image is of 145 by 145 pixels, with a spatial resolution of 20 m/pixel and 200 spectral channels. A three-band false color image and the reference data are shown in Fig. 2. Sixteen classes of interest are considered, which are detailed in Table II, with a number of training and test samples for each class. Training samples have been randomly chosen from the reference data.
- 2) The *Center of Pavia* image was recorded by the ROSIS sensor over the urban area of Pavia, Italy. It is of 900 by 300 pixels, with a spatial resolution of 1.3 m/pixel and 102 spectral channels. The reference data contain nine thematic classes and 56 070 labeled pixels. Thirty samples for each class were randomly chosen from the reference data as training samples.
- 3) The *University of Pavia* image is of an urban area, acquired by the ROSIS sensor. It is of 610 by 340 pixels, with 103 spectral channels. The reference data contain nine classes of interest. The training and test sets are composed of 3921 and 40 002 pixels, respectively.

More information about the images can be found in [8].

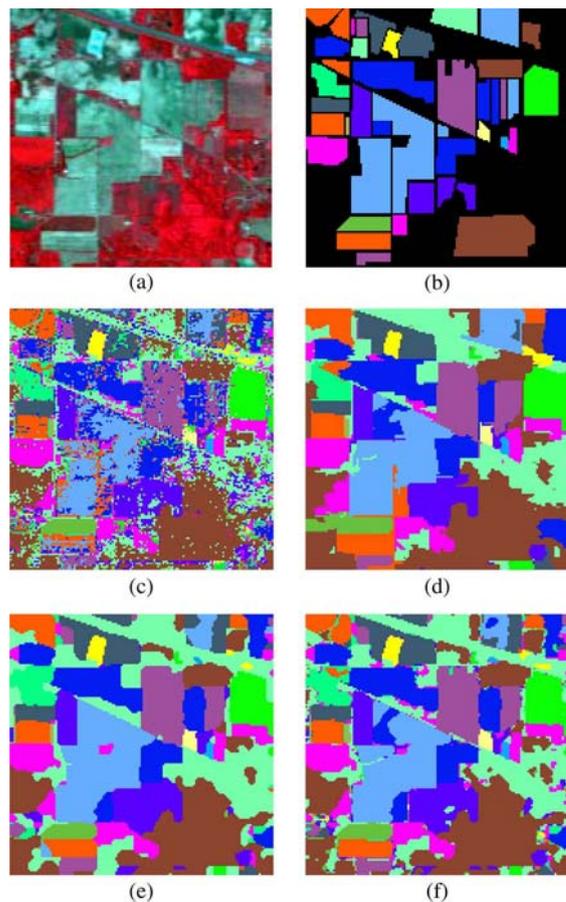


Fig. 2. *Indian Pines* image. (a) Three-band color composite. (b) Reference data. (c) SVM pixelwise classification map. (d) SVMMSF + MV classification map. (e) SVMRF-NE classification map. (f) SVMRF-E classification map.

In all experiments, the probabilistic one-versus-one SVM classification with the Gaussian radial basis function (RBF) kernel was applied. The optimal parameters C (parameter that controls the amount of penalty during the SVM optimization [4]) and γ (spread of the RBF kernel) were chosen by fivefold cross validation. The temperature T was varied during the Metropolis relaxation procedure [18]: The initial temperature was set to $T^1 = 2$ (a relatively low value of the initial temperature results in a faster execution of the algorithm). After every 10^6 (order of the number of pixels in an image) iterations, the temperature for the next iteration ($k + 1$) was recomputed as $T^{k+1} = 0.98T^k$. The optimal value of the parameter $\alpha = 30$ was experimentally derived (the same optimal value of α was obtained for the three considered data sets).

Furthermore, we have investigated the performances of the SVMRF-NE and SVMRF-E algorithms for different values of the context weight parameter β . Table I reports the SVMRF-NE and SVMRF-E overall (percentage of correctly classified pixels) and average (mean of the percentage of correctly classified pixels for each class) classification accuracies for the three considered data sets. It can be concluded that the optimal parameter is $\beta \in [1, 2]$ for the SVMRF-NE approach and $\beta \in [2, 4]$ for the SVMRF-E approach (for both methods, the corresponding overall accuracies are nonsignificantly different over the given range of values). Moreover, the

TABLE I
SVMMRF-NE AND SVMMRF-E CLASSIFICATION ACCURACIES FOR DIFFERENT VALUES OF THE PARAMETER β

Image	Accuracy, %	β for SVMMRF-NE					β for SVMMRF-E						
		0.5	1	2	3	5	0.5	1.0	2.0	3.0	4.0	5.0	10.0
Indian Pines	Overall	90.42	92.05	92.31	92.53	92.78	84.99	88.19	91.83	91.59	90.64	90.03	90.23
	Average	94.70	95.83	89.81	83.24	91.76	90.50	92.85	95.69	95.37	92.62	95.01	88.67
Center of Pavia	Overall	97.26	97.38	97.26	97.31	97.27	97.11	97.46	97.56	97.60	97.50	97.44	97.33
	Average	93.98	94.13	93.94	94.05	93.73	93.41	94.19	94.24	94.59	94.23	94.12	93.98
University of Pavia	Overall	85.30	86.40	86.89	87.57	86.93	84.20	85.79	87.18	87.63	87.57	87.63	87.50
	Average	91.64	91.72	92.12	91.90	91.75	91.03	92.22	92.97	93.41	93.27	93.28	93.39

TABLE II
NUMBER OF LABELED SAMPLES (NUMBER OF SAMPLES) AND CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE *INDIAN PINES* IMAGE

	No. of Samp.		SVM	ECHO	WH+MV	SVMMSF+MV	SVMMRF-NE	SVMMRF-E
	Train	Test						
Overall Accuracy	-	-	78.17	82.64	86.63	91.80	92.05	91.83
Average Accuracy	-	-	85.97	83.75	91.61	94.28	95.83	95.69
Kappa Coefficient (κ)	-	-	75.33	80.38	84.83	90.64	90.93	90.71
Corn-no till	50	1384	78.18	83.45	94.22	93.21	93.28	98.48
Corn-min till	50	784	69.64	75.13	78.06	96.56	83.93	90.82
Corn	50	184	91.85	92.39	88.59	95.65	99.46	98.37
Soybeans-no till	50	918	82.03	90.10	96.30	93.91	98.58	98.91
Soybeans-min till	50	2418	58.95	64.14	68.82	81.97	82.09	76.92
Soybeans-clean till	50	564	87.94	89.89	90.78	97.16	97.70	97.34
Alfalfa	15	39	74.36	48.72	94.87	94.87	97.44	97.44
Grass/pasture	50	447	92.17	94.18	95.08	94.63	97.54	97.54
Grass/trees	50	697	91.68	96.27	97.99	97.27	97.70	97.56
Grass/pasture-mowed	15	11	100	36.36	100	100	100	100
Hay-windrowed	50	439	97.72	97.72	99.54	99.77	99.54	99.54
Oats	15	5	100	100	100	100	100	100
Wheat	50	162	98.77	98.15	99.38	99.38	99.38	99.38
Woods	50	1244	93.01	94.21	97.11	99.68	98.39	99.04
Bldg-Grass-Tree-Drives	50	330	61.52	81.52	69.39	68.79	88.18	79.70
Stone-steel towers	50	45	97.78	97.78	95.56	95.56	100	100

methods are robust to the choice of β , and quite a wide range of values of β leads to high classification accuracies.¹

Table II summarizes the global (overall average accuracies and kappa coefficient [8]) and class-specific classification accuracies for the *Indian Pines* image. In order to compare the performances of the proposed method with other recently proposed advanced classification techniques, we have included results of the pixelwise SVM classifiers, the well-known ECHO (Extraction and Classification of Homogeneous Object) spatial classifier [5], classification using majority vote within the adaptive neighborhoods defined by watershed segmentation (WH + MV) [19], as well as the results obtained using the construction of a minimum spanning forest from the probabilistic SVM-derived markers followed by majority voting within connected regions (SVMMSF + MV) [8]. Fig. 2 shows some of the corresponding classification maps. As can be seen from the table, all the spectral-spatial approaches yield higher classification accuracies when compared to the pixelwise method. The proposed SVMMRF-NE and SVMMRF-E techniques give the highest global and most of the best class-specific accuracies. Following the results of the McNemar's test, the SVMMRF-NE, SVMMRF-E, and SVMMSF + MV accuracies are not significantly different, using 5% level of significance. From Fig. 2, it can be seen that the corresponding three classification maps are comparable and contain more homogeneous regions, when compared to the SVM classification map. Since the considered

TABLE III
GLOBAL CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE *CENTER OF PAVIA* IMAGE

	Overall Accuracy	Average Accuracy	Kappa Coefficient (κ)
SVM	95.64	90.60	92.71
ECHO	96.22	92.47	93.70
WH+MV	96.26	92.08	93.75
SVMMSF+MV	96.62	92.78	94.35
SVMMRF-NE	97.38	94.13	95.61
SVMMRF-E	97.60	94.59	95.97

image contains large spatial structures and reference data do not comprise region edges, the advantage of the SVMMRF-E method versus the SVMMRF-NE method is not obvious here.

Table III gives the global classification accuracies for the *Center of Pavia* data, where the same techniques are used for comparison. The proposed SVMMRF-E method yields the best classification accuracies. This image of an urban area contains small spatial structures, such as shadows and trees. Therefore, the inclusion of the edge information in the context-based regularization improves the classification performances.

Table IV reports the global classification accuracies for the *University of Pavia* image. For this data set, the SVMMSF + MV classifier gives the best accuracies, and the SVMMRF-E method outperforms the SVMMRF-NE technique in terms of accuracies. According to the results of the McNemar's test, all the corresponding classification maps are significantly different, using 5% level of significance. From these results, the following conclusions can be derived: 1) the advantage of the

¹A similar study has shown robustness of the SVMMRF-E method to the choice of the parameter α .

TABLE IV
GLOBAL CLASSIFICATION ACCURACIES IN PERCENTAGE
FOR THE UNIVERSITY OF PAVIA IMAGE

	Overall Accuracy	Average Accuracy	Kappa Coefficient (κ)
SVM	81.01	88.25	75.86
ECHO	87.58	92.16	83.90
WH+MV	85.42	91.31	81.30
SVMMMSF+MV	91.08	94.76	88.30
SVMMRF-NE	86.89	92.12	83.14
SVMMRF-E	87.63	93.41	84.07

edge-based SVMMRF-E method for the classification of urban images is confirmed, and 2) the MRF-based regularization includes the spatial context information from only the closest neighborhoods (in our case, eight neighborhoods) when classifying an image. Therefore, the proposed method is efficient only in the case if there is no large misclassified region in the initial pixelwise classification map (this assumption often holds). If such a region exists, the MRF-based method cannot reconstitute its true class label. This happens in the case of classification of the *University of Pavia* image, where some relatively large regions remain misclassified. The SVMMMSF + MV method works differently: If there is a suspicion that a pixel might be misclassified, this pixel remains unclassified, and the classification decision is further taken by the region growing step. This appears to be a robust procedure for classifying large regions with uncertain spectral properties.

Furthermore, the proposed SVMMRF-E method was applied and has shown the best classification (overall and average) accuracies at the Third HYPER-I-NET summer school on hyperspectral imaging student contest "Evaluation of an unknown hyperspectral data set and information extraction." The contest was organized by P. Gamba on September 10, 2009, in Pavia, Italy, where the students in teams were supposed to provide a classification map of the rural area acquired by the Compact Airborne Spectrographic Imager sensor.

When comparing the results of several classifiers, an important issue is the computational cost of each classifier. Here, we compare the computational times for classification of the *University of Pavia* image using different methods. We conducted experiments on an Intel Core 2 Duo 2.40-GHz processor with 3.5-GB RAM. The processing times in seconds were 3339 for the SVM method, 3353 for the WH + MV method, 3351 for the SVMMMSF + MV method, 3444 for the SVMMRF-NE method, and 3450 for the SVMMRF-E method. None of the algorithms has been implemented in parallel (which would further speed up computational times). While the SVM classifier is a computationally demanding algorithm, other considered methods require at maximum 3% more time to be executed. In terms of duration, the proposed SVMMRF-NE method takes 93 s longer for the classification of the data than the SVMMMSF + MV approach, and the SVMMRF-E method takes 6 s longer than the SVMMRF-NE method.

IV. CONCLUSION

A novel accurate SVMMRF method for spectral-spatial classification of hyperspectral images has been presented in this

letter. The method consists in performing a probabilistic SVM pixelwise classification, followed by MRF-based regularization for incorporating spatial and edge information into classification. Experimental results have demonstrated that the proposed method yields accurate classification maps within a short time interval and is sufficiently robust for classifying different kinds of images.

ACKNOWLEDGMENT

The authors would like to thank D. Landgrebe from Purdue University, and P. Gamba from the University of Pavia, for providing the hyperspectral data.

REFERENCES

- [1] C.-I Chang, *Hyperspectral Data Exploitation: Theory and Applications*. New York: Wiley-Interscience, 2007.
- [2] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. New York: Wiley, 2003.
- [3] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [4] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
- [5] R. L. Kettig and D. A. Landgrebe, "Classification of multispectral image data by extraction and classification of homogeneous objects," *IEEE Trans. Geosci. Electron.*, vol. GE-14, no. 1, pp. 19–26, Jan. 1976.
- [6] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.
- [7] A. Farag, R. Mohamed, and A. El-Baz, "A unified framework for map estimation in remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, Jul. 2005.
- [8] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, 2010, to be published. [Online]. Available: DOI: 10.1109/TSMCB.2009.2037132.
- [9] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [10] F. Bovolo and L. Bruzzone, "A context-sensitive technique based on support vector machines for image classification," in *Proc. PRMI*, 2005, pp. 260–265.
- [11] D. Liu, M. Kelly, and P. Gong, "A spatial-temporal approach to monitoring forest disease spread using multi-temporal high spatial resolution imagery," *Remote Sens. Environ.*, vol. 101, no. 2, pp. 167–180, Mar. 2006.
- [12] A. Solberg, T. Taxt, and A. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.
- [13] J. Platt, "Probabilities for support vector machines," in *Advances in Large Margin Classifiers*, A. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans, Eds. Cambridge, MA: MIT Press, 2000, pp. 61–74.
- [14] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multiclass classification by pairwise coupling," *J. Mach. Learn. Res.*, no. 5, pp. 975–1005, 2004.
- [15] Y. Tarabalka, J. Chanussot, and J. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognit.*, vol. 43, no. 7, pp. 2367–2379, Jul. 2010.
- [16] R. Gonzalez and R. Woods, *Digital Image Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [18] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [19] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in *Proc. IGARSS*, Boston, MA, 2008, pp. III-652–III-655.

Paper 2

Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, “Segmentation and classification of hyperspectral images using watershed transformation,” *Pattern Recognition*, 2010, vol. 43, no. 7, pp. 2367-2379, July 2010.

Segmentation and classification of hyperspectral images using watershed transformation

Y. Tarabalka^{a,b,*}, J. Chanussot^a, J.A. Benediktsson^b

^a GIPSA-Lab—Grenoble Institute of Technology, Domaine Universitaire—BP 46—38402 Saint-Martin-d'Hères Cedex, France

^b Faculty of Electrical and Computer Engineering — University of Iceland, Hjarðarhaga 2-6, 107 Reykjavík, Iceland

ARTICLE INFO

Article history:

Received 17 March 2009

Received in revised form

13 November 2009

Accepted 21 January 2010

Keywords:

Hyperspectral images

Mathematical morphology

Watershed

Segmentation

Classification

ABSTRACT

Hyperspectral imaging, which records a detailed spectrum of light for each pixel, provides an invaluable source of information regarding the physical nature of the different materials, leading to the potential of a more accurate classification. However, high dimensionality of hyperspectral data, usually coupled with limited reference data available, limits the performances of supervised classification techniques. The commonly used pixel-wise classification lacks information about spatial structures of the image. In order to increase classification performances, integration of spatial information into the classification process is needed. In this paper, we propose to extend the watershed segmentation algorithm for hyperspectral images, in order to define information about spatial structures. In particular, several approaches to compute a one-band gradient function from hyperspectral images are proposed and investigated. The accuracy of the watershed algorithms is demonstrated by the further incorporation of the segmentation maps into a classifier. A new spectral-spatial classification scheme for hyperspectral images is proposed, based on the pixel-wise Support Vector Machines classification, followed by majority voting within the watershed regions. Experimental segmentation and classification results are presented on two hyperspectral images. It is shown in experiments that when the number of spectral bands increases, the feature extraction and the use of multidimensional gradients appear to be preferable to the use of vectorial gradients. The integration of the spatial information from the watershed segmentation in the hyperspectral image classifier improves the classification accuracies and provides classification maps with more homogeneous regions, compared to pixel-wise classification and previously proposed spectral-spatial classification techniques. The developed method is especially suitable for classifying images with large spatial structures.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The growing availability of hyperspectral images has opened the door to numerous new applications in remote sensing and other areas of image analysis. Hyperspectral sensors capture more than a hundred spectral bands (data channels) simultaneously. Thus, each pixel in a hyperspectral image is presented as the vector of values corresponding to the wide spectrum of reflected light [1] (Fig. 1 depicts the structure of a hyperspectral image). For instance, NASA Jet Propulsion Laboratory's Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) system has 224 spectral channels with a spectral resolution of around 10 nm, covering the wavelengths from

0.4 to 2.5 μm [2]. The fine spectral resolution of the data provides an invaluable source of information regarding the physical nature of the different materials, increasing the capability to distinguish structures and objects in the image scene.

However, such a large number of spectral channels implies the high dimensionality of the data and presents challenges to image analysis and classification. Most of the commonly used methods designed for the analysis of grey level, color or multispectral images are not appropriate for hyperspectral images. As a matter of fact, very limited reference data are usually available (the ratio of the number of referenced samples to the number of spectral channels quickly drops), which limits the performances of supervised classification techniques. Furthermore, for analysis of hyperspectral images the well known curse of dimensionality prevents robust statistical estimations, usual vector norms become meaningless and so on (e.g., the Hughes phenomenon [3]). Therefore, to take full advantage of the rich information provided by the spectral dimension, the development of new algorithms is required.

* Corresponding author at: GIPSA-Lab—Grenoble Institute of Technology, Domaine Universitaire—BP 46—38402 Saint-Martin-d'Hères Cedex, France. Tel.: +33 6 66 95 72 75; fax: +33 4 76 82 63 84.

E-mail addresses: yuliya.tarabalka@hyperinet.eu (Y. Tarabalka), jocelyn.chanussot@gipsa-lab.grenoble-inp.fr (J. Chanussot), benedikt@hi.is (J.A. Benediktsson).

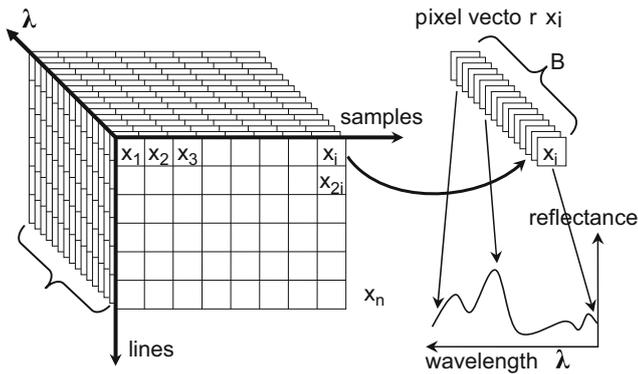


Fig. 1. Structure of a hyperspectral image.

The first attempts to classify hyperspectral images were designed to assign each pixel to one of the classes based on its spectrum only [4]. These *pixel-level* processing systems use a wide range of features, such as the direct spectral information, texture features, and linear and nonlinear transformations of these features. The applied feature extraction procedure often aims at reducing the dimensionality of the data. The features are used for image classification with a wide range of techniques, such as maximum-likelihood or Bayesian estimation techniques [5–7], neural networks [8–10], decision trees [11,12], genetic algorithms [13] and kernel-based methods [14–17]. In particular, Support Vector Machines (SVM) and other kernel-based methods have recently shown good classification results, because they tend to be robust when a limited number of training samples is available.

Further modification to improve classification results consists in the integration of spatial and spectral information in the image analysis. It means that the decision to assign a pixel to a specific class is simultaneously based on the feature vector of this pixel and on some information derived from the pixel's neighborhood. In previous studies, the spatial context was included in the classification system by using morphological filters [15], morphological leveling [18] and Markov random fields [19]. These methods show promising results in terms of classification accuracies by incorporating spatial and spectral information. However, these algorithms use fixed-window-based neighborhoods. This involves the problem of scale selection, especially if the image contains some small or complex structures.

Another approach to define spatial structures consists in performing image segmentation [20–22]. The regions obtained from the realized segmentation map define the spatial context of the pixels within these regions. To make this approach effective, an accurate segmentation of the image is needed.

In previous studies, several methods for multispectral image segmentation have been investigated. Numerous works exploit region merging techniques, where neighboring image segments are merged iteratively based mostly on their spectral similarity. For instance, the eCognition software performs multiresolution segmentation, based on bottom-up region merging [23]. Initially, each pixel is considered as a separate region, and subsequently pairs of regions are merged, based on a homogeneity criterion, which is a combination of spectral and shape properties. Tilton developed a hierarchical segmentation algorithm [24], which performs region growing and spectral clustering alternately. The main drawback of applying region merging for image segmentation is that the homogeneity criterion, or thresholds must be chosen. For accurate segmentation, these techniques usually produce a pyramid of segmentation maps, using a range of thresholds. Then, manual interpretation of the results is needed.

Other studies exploit mathematical morphology based segmentation approaches [18,25–30], which mostly use

granulometries or watershed transformation. The extension of morphological operators to the case of multispectral images is not straightforward, because there is no natural way for total ordering of multivariate pixels, which is a requirement in mathematical morphology. An extensive literature on mathematical morphology for color and multispectral images is available [31–36]. In particular, the watershed segmentation of color images was investigated in [37–39].

However, the above morphological and watershed methods are not suitable for segmentation of hyperspectral images, due to the following reasons:

- A hyperspectral image is composed of hundreds of spectral channels. Therefore, the use of total ordering schemes for multivariate data, such as the bit mixing paradigm [31], is not possible, because it would lead to a huge number of values stored for each pixel.
- In previous studies, polar-based representations (HLS, HSV) and perceptual color spaces (LUV, LAB) were used for morphological analysis of color images [40,37]. These methods are not applicable for hyperspectral images.

In a recent paper, Noyel et al. [29] gave an overview of the literature on the watershed-based multispectral image segmentation and performed watershed segmentation of hyperspectral images. Their method is composed of spectral classification to obtain markers and computation of a multivariate gradient to get spatial information. Only visual results (the obtained segmentation maps) are presented in the article. Therefore, the question of defining a watershed transformation for the case of hyperspectral images has only recently been raised in the literature and needs further investigations.

Some studies have been conducted on spectral-spatial classification of multispectral images. Linden et al. [41] used the vector mean as a feature for each region in a spectral-spatial classifier. First, they performed the segmentation of a hyperspectral image, based on region growing (using the eCognition software [23]). Then, a vector mean was computed for each region, such that the value in each band represented the average spectral information of the pixels in this region in the respective band. Afterwards, the regions were classified by an SVM classifier. However, the obtained results were mostly not an improvement over those obtained by the pixel-wise SVM classification. Li and Xiao [28] used spectral and spatial information for classification of a multispectral (4-band SPOT 5) image. There, a watershed segmentation and a pixel-wise maximum likelihood classification of an image were independently performed. Then, pixels of the whole region were assigned to one class if more than 50% of pixels in this region were categorized into one class by a pixel-wise classifier. The classification results were substantially improved with the spectral-spatial approach compared to the pixel-wise maximum likelihood classification.

Widayati et al. [42] used spatial information in a classifier, in order to perform a spectral-spatial classification of a multispectral (4-band IKONOS) image. First, a segmentation map was obtained using the *Merge Using Moments* algorithm [43]. Then two options were explored. First, each region from the segmentation map was classified using its vector mean as a feature. In another approach, the pixel-wise classification map obtained by maximum likelihood classifier was combined with the segmentation map using majority voting: for every region from the segmentation map, all the pixels were assigned to the majority class within this region. Results of the pixel-wise maximum likelihood classification were also used for comparison. Of all these three

methods, the spectral–spatial classification using majority voting gave the highest overall accuracy.

The main contributions of this paper are two-fold:

1. *The first contribution* is the extension of the watershed segmentation algorithm to hyperspectral data. Watershed transformation is usually applied to the gradient, which must be a scalar function, in order to form a complete lattice structure [44]. When processing a hyperspectral image, information about spatial structures must be extracted from all bands in the optimal way. Different approaches to segment the hyperspectral image by watershed are proposed and investigated. In particular, different ways to compute a one-band gradient function of a hyperspectral image are considered. We emphasize that this study is not a comparison of watershed segmentation methods in general. Instead, it focuses on the extension of watershed transformation for hyperspectral images, for the purpose of accurate segmentation and further classification.
2. The obtained watershed segmentation map is further incorporated into a spectral–spatial classifier, aiming at improving classification accuracies, when compared to pixel-wise classification. Thus, *the second contribution* of this paper is the development of the new segmentation and classification scheme to analyze hyperspectral data. The proposed classification method combines results of a pixel-wise SVM classification and the segmentation map using majority vote approach [21]. This contribution can be summarized as follows: the segmentation defines an adaptive neighborhood for each pixel which is used for the spatial regularization following a pixel-wise classification.

To test the developed segmentation and classification algorithms, two hyperspectral airborne images are used: A 103-band ROSIS (Reflective Optics System Imaging Spectrometer) image of the University of Pavia, Italy, and a 220-band AVIRIS image taken over the Northwestern Indiana's Indian Pine site [45].

The paper is organized as follows. In Section 2, an overview of the watershed technique is given, and then the extension of the watershed algorithm to hyperspectral images is discussed. In Section 3, the developed segmentation and classification scheme is presented. In Section 4, experimental results and comparisons are presented and discussed. Finally, conclusions are drawn in Section 5.

2. Watershed segmentation

Watershed transformation is a powerful mathematical morphology technique for image segmentation [46,44]. It was introduced in image analysis by Beucher and Lantuéjoul [47], and subsequently a lot of algorithms for its implementation have been proposed.

The watershed transform considers a two-dimensional one-band image as a topographic relief. The value of a pixel h stands for its elevation. The watershed lines divide the image into catchment basins, so that each basin is associated with a *minimum* in the image (see Fig. 2). The watershed transformation is usually applied to the gradient function of the image. The gradient defines transitions between regions, so that it has high values on the borders between objects and *minima* in the homogeneous regions. And in this case, if the crest lines in the gradient image correspond to the edges of image objects, watershed transformation partitions this image into meaningful regions.

A wealth of literature describes techniques for computing the watershed transformation (see for instance [48–51]). A review of

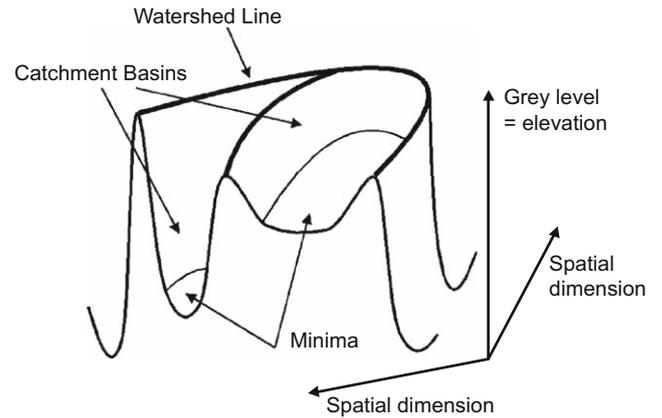


Fig. 2. Topographic representation of a one-band image.

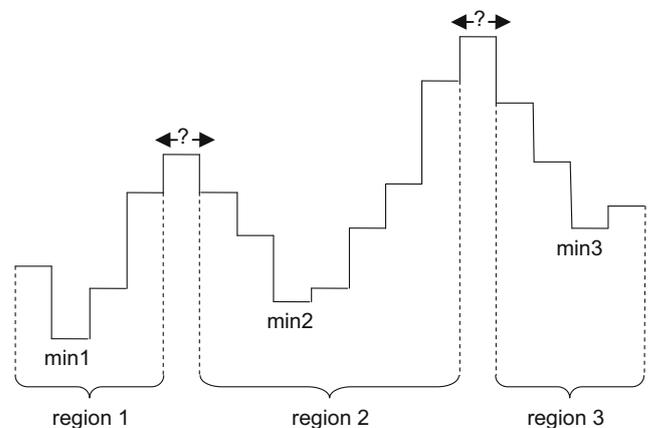


Fig. 3. Example of watershed transformation in one dimension.

watershed algorithms can be found in [48,50]. Vincent and Soille [48] have proposed an efficient watershed algorithm using flooding simulations, which has become one of the classical algorithms to compute watersheds.

The output of the watershed transform is a partition of the image composed of regions (sets of pixels connected to the same local *minimum*) and of *watershed pixels* (WHEDs, the borders between the regions). Fig. 3 shows an example of watershed transformation in one dimension, where three regions, associated with the three *minima*, are defined. The two *maxima* correspond to the borders between regions and are not assigned to any region (watershed pixels).

Typically, the result of watershed segmentation on the gradient image without any additional processing is a severe oversegmentation (every single local *minimum* of the gradient leads to one region). Common ways to reduce the number of local *minima* are to filter the original image or the gradient function (e.g., area filtering) or to use markers [46]. The oversegmentation effect can be also corrected using some post-processing, such as merging of similar neighboring regions.

In the next subsection, different approaches for segmentation of hyperspectral images by watershed are discussed and extended.

2.1. Watershed segmentation of hyperspectral images

As previously mentioned, the watershed transformation requires as input a one-band image and gives as a result a

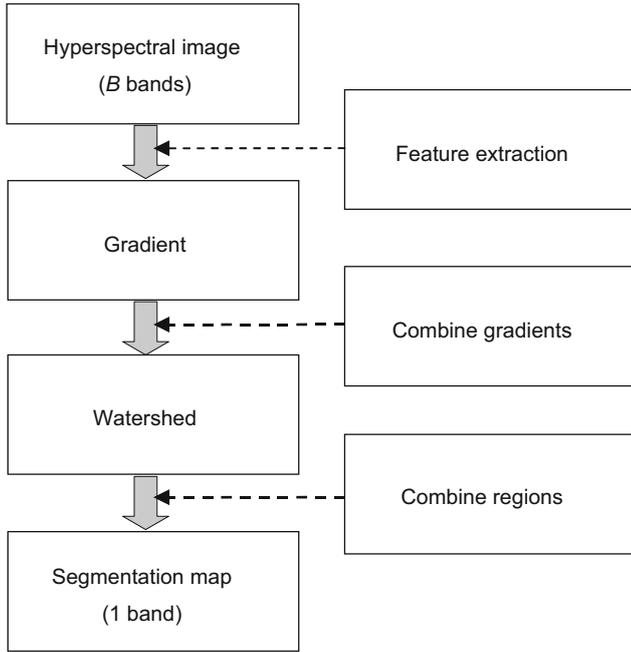


Fig. 4. Flow-chart which shows strategies of applying watershed to a hyperspectral image.

one-band segmentation map (where each pixel contains the label of the catchment basin or the watershed pixel label). In this paper, we aim to apply this transformation to a B -band hyperspectral image. Let us consider this image as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$ (each pixel is characterized by its spatial location and a vector of spectral values; see Fig. 1). We also denote the image of every spectral band as $X_\lambda, \lambda = 1, 2, \dots, B$. Different strategies are possible to compute watersheds. They are summarized in Fig. 4.

Before computing a gradient, *feature extraction* on the original image can be performed, applying one of the transformations such as the principal component analysis (PCA) [52,53], maximum noise fraction (MNF) [54] and independent component analysis (ICA) [55,56]. The aim of this step is to obtain either a one-band image or a multi-band image which would contain enough information to distinguish between spatial structures in the image.

If a one-band image with a good distinguishing capability between structures can be obtained, the algorithm for computing a gradient and watershed is straightforward. For instance, for a one-band image Y , a basic morphological gradient (also called the Beucher gradient) can be applied. It is defined as the arithmetic difference between the dilation δ and the erosion ε of Y by the structuring element E [44]:

$$\rho_E(Y) = \delta_E(Y) - \varepsilon_E(Y). \quad (1)$$

If, however, at the input of the *gradient* step we still have a multi-band image, we can proceed in different ways, that can be grouped into the following three categories:

- to compute a vectorial gradient;
- to compute a multidimensional gradient;
- to combine watershed segmentation maps *a posteriori*.

These three options are discussed in the next three subsections.

2.1.1. Computation of a vectorial gradient

Vectorial gradients are based on the distance between pixel vectors, and produce from the B -band image one-band gradient

[57,29]. Several types of vectorial gradients have been proposed. Noyel et al. [29] proposed to use a metric-based gradient for hyperspectral images, which is described as follows: For each pixel vector \mathbf{x}_p , let $\psi = [\mathbf{x}_p^1, \mathbf{x}_p^2, \dots, \mathbf{x}_p^e]$ be a set of e vectors in the neighborhood of \mathbf{x}_p (set ψ does not contain \mathbf{x}_p). For instance, a four- or an eight-neighborhood ($e = 4$ or $e = 8$, respectively) can be used. The *metric-based gradient* is defined as a difference between the *supremum* and the *infimum* of the defined distances between \mathbf{x}_p and vectors from the set ψ :

$$\nabla_{\psi,d}^{MB}(\mathbf{x}_p) = \sup_{i \in \psi} \{d(\mathbf{x}_p, \mathbf{x}_p^i)\} - \inf_{j \in \psi} \{d(\mathbf{x}_p, \mathbf{x}_p^j)\}. \quad (2)$$

Various distances can be used to compute gradient from (2) such as Euclidean, Mahalabobis, chi-squared distances [29].

Another type of vectorial gradient is the *Robust Color Morphological Gradient* (RCMG). This gradient was developed for color images by Evans and Liu [57]. Here we investigate the use of the RCMG for hyperspectral images.

For each pixel vector \mathbf{x}_p , let $\chi = [\mathbf{x}_p^1, \mathbf{x}_p^2, \dots, \mathbf{x}_p^e]$ be a set of e vectors within a structuring element E , which defines the neighborhood of the vector \mathbf{x}_p , and the set χ contains \mathbf{x}_p . The *Color Morphological Gradient* (CMG) is computed as

$$\nabla_{\chi,d}^{CM}(\mathbf{x}_p) = \max_{i,j \in \chi} \{d(\mathbf{x}_p^i, \mathbf{x}_p^j)\}, \quad (3)$$

i.e., the *maximum* of the distances between all pairs of vectors in the set χ . Here, various distances can be chosen. If the Euclidean distance is used, (3) can be rewritten as

$$\nabla_{\chi,d}^{CM}(\mathbf{x}_p) = \max_{i,j \in \chi} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}. \quad (4)$$

One of the drawbacks of the CMG is that it is very sensitive to noise. To overcome the problem of outliers, the authors of [57] have proposed to use the RCMG. The scheme to make the CMG robust consists in removing the two pixels that are the furthest apart and then finding the CMG of the remaining pixels. This process can be repeated several times until a good estimate of the gradient is obtained. Thus, the *RCMG*, using the Euclidean distance, can be defined as

$$\nabla_{\chi,d}^{RCM}(\mathbf{x}_p) = \max_{i,j \in \{\chi - REM_r\}} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}, \quad (5)$$

where REM_r is the set of the r vector pairs removed. The appropriate value of the parameter r in (5) depends on the chosen structuring element E and the amount of noise present in the image, as discussed in [57]. When a one-band vectorial gradient is computed, it is used as the input of the watershed algorithm.

2.1.2. Multidimensional gradient methods

Another approach to compute a one-band gradient from the multi-band image consists in considering the B -band image as a set of B one-band images. In this case, the gradients of every spectral band can be computed, using for instance a morphological gradient (see (1)). Then the obtained B gradient images $\rho_E(X_\lambda), \lambda = 1, 2, \dots, B$ are combined into one image using linear or nonlinear operators.

As an example of the linear operators, the weighted sum of gradients can be computed by

$$\nabla_E^+(\mathbf{X}) = \sum_{\lambda=1}^B \omega_\lambda \rho_E(X_\lambda), \quad (6)$$

where ω_λ denotes the weight of the gradient of the band λ . If $\omega_\lambda = 1, \lambda = 1, 2, \dots, B$, all the bands are supposed to have an equal importance in defining the gradient. Modifying the weight coefficients, the gradient estimation can be improved. For instance, coefficients that are inversely proportional to the

estimated noise of each spectral band can be used as the weights in (6).

Examples of nonlinear operators are the *supremum* and the median operators. The gradient *supremum* over morphological gradients of every band is defined as follows: for every pixel the *supremum* over all gradient images is taken as the output value of this pixel.

2.1.3. Combination of watershed segmentation maps

First, B standard gradients are computed, one for each spectral band. Then each gradient image is used to compute a watershed transformation. This gives B segmentation maps that can be further combined to provide a single output segmentation map.

One of the ways to combine the B segmentation maps, in order to define relevant edges, consists in summing the watershed lines. Here, for each segmentation map obtained from the gradient of band λ , a binary image W_λ of watershed lines is produced. Thus, W_λ is an image in which watershed pixels are equal to 1 and all other pixels are equal to 0. The sum of watershed lines is computed by

$$\mathbf{W} = \sum_{\lambda=1}^B W_\lambda. \quad (7)$$

The obtained image \mathbf{W} can be further thresholded, in order to define the border pixels that were presented in most of the segmentation maps, hence ensuring a reliable edge detection.

However, when summing the watershed lines, we do not have information about regions anymore, but only about edges. Furthermore, some edges can become open after thresholding. Therefore, closing of edges and image region labeling must be performed after the procedure described above.

3. Segmentation and classification scheme

As previously mentioned, the information about spatial structures defined by the watershed segmentation algorithm can be used to improve the results of classification of a hyperspectral image. In this section, a new combined spectral-spatial classification scheme is presented for hyperspectral images based on watershed segmentation.

The general flow-chart of the proposed segmentation and classification scheme is given in Fig. 5. At the input we have a B -band hyperspectral image $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. First, the watershed segmentation is performed on this image, using one of the approaches discussed in Section 2 and summarized in Fig. 4. In the resulting segmentation map, each pixel contains the label of the region it belongs to, or the watershed pixel label (see Figs. 3 and 6).

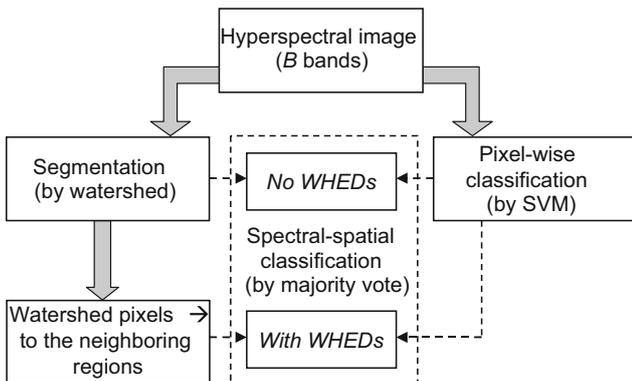


Fig. 5. Flow-chart of the proposed segmentation and classification scheme.

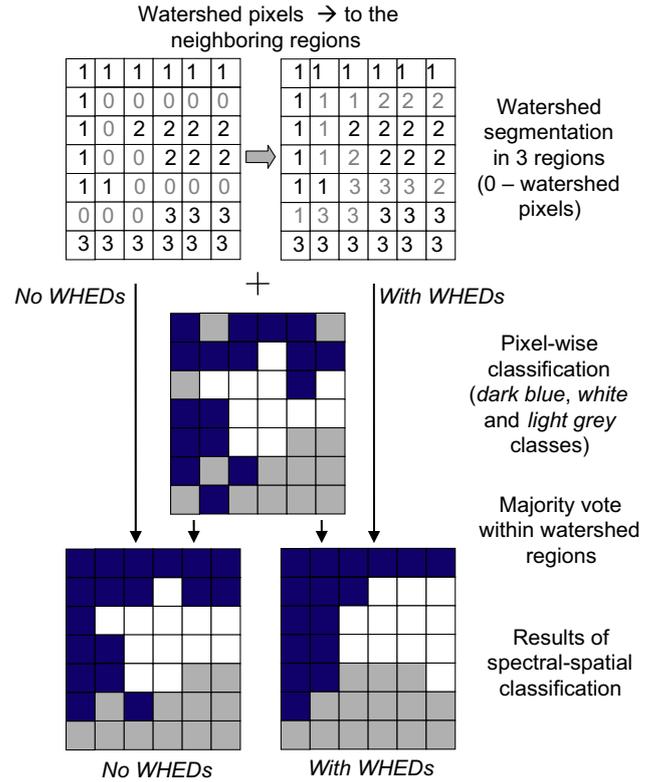


Fig. 6. Example of spectral-spatial classification.

It is often desirable to produce a segmented image where each pixel belongs to some region, without border pixels between regions. In this case, each watershed pixel can be assigned to one of the regions in its neighborhood. For this purpose, we propose to compute for every region \mathbf{S} ($\mathbf{S} = \{\mathbf{s}_j \in \mathbb{R}^B, j = 1, 2, \dots, l\}$, $\mathbf{S} \subseteq \mathbf{X}$, with l equal to the number of pixels in the region) the standard vector median [58]. A standard vector median \mathbf{s}_{VM} for a set of pixel vectors is a vector, which fulfills the condition that the sum of the distances between this vector and all the other vectors in the set is minimal (for instance, L_1 norm is used to compute distances):

$$\mathbf{s}_{VM} = \operatorname{argmin}_{\mathbf{s} \in \mathbf{S}} \left\{ \sum_{j=1}^l \|\mathbf{s} - \mathbf{s}_j\|_1 \right\}. \quad (8)$$

Every watershed pixel is assigned to the neighboring region with the "closest" median, i.e., the distance between the vector median of this region and the watershed pixel vector is minimal (see an example in Fig. 6).

After the image is segmented into regions, this spatial information should be used to improve the classification results. Two approaches to integrate spectral and spatial information into the classification system can be distinguished:

1. To define a feature or a set of features for each region from the segmentation map, and classify regions using these features.
2. To perform a pixel-wise classification first, and then combine a pixel-based classified image with the segmentation results.

Here we propose to use the second approach for the spectral-spatial classification of hyperspectral images, in order to improve the results of the pixel-wise classifier using the spatial information from the segmentation map (see Figs. 5 and 6). The proposed approach is outlined in [21]. First, a pixel-wise classification by SVM on the original hyperspectral image is performed. Then, for every watershed region \mathbf{S} , all the pixels are assigned to the most

frequent class within this region (*majority vote* approach). Please note that majority vote within fixed neighborhoods is a standard spatial regularization procedure following a pixel-wise classification. Here, we propose to use the results of a segmentation to define an adaptive neighborhood for each pixel.

The watershed pixels can be either left not processed during the majority voting (what we call the *No WHEDs* approach), or assigned to the regions with the “closest median” before the majority voting is performed considering all the pixels (what we call the *With WHEDs* approach). Fig. 6 shows an example of the combination of the pixel-wise classification map and the segmentation map by majority vote, using both *No WHEDs* and *With WHEDs* approaches.

4. Experimental results and discussion

4.1. Segmentation and classification of the University of Pavia image

4.1.1. Dataset

The *University of Pavia* image is of an urban area that was recorded by the ROSIS-03 optical sensor. The image scene is the

urban area surrounding the University of Pavia, Italy. The image has spatial dimensions of 610 by 340 pixels, with a spatial resolution of 1.3 m per pixel. The number of bands of the ROSIS-03 sensor is 115 with a spectral coverage ranging from 0.43 to 0.86 μm . The 12 most noisy channels have been removed, and the experiments are conducted on the 103-band image. Nine classes of interest are considered, with the number of test and training samples detailed for each class in Table 1. Training-test set was provided by Prof. Paolo Gamba together with the hyperspectral image. False color image of the University of Pavia and the reference data are presented in Fig. 7.

4.1.2. Experimental results

Different approaches of the watershed transformation of hyperspectral images described in Section 2 were tested. Watershed segmentation was performed on the gradient function obtained in four different ways:

1. *Band50*: A morphological Beucher gradient (see (1)) was computed on one band. Band no. 50 was chosen arbitrarily, but similar results were obtained with other non-noisy bands. Here and in all the following morphological operations, a 3×3

Table 1
Information classes, training-test samples, and class-specific accuracies in percentage for the *University of Pavia* image.

Class	Samples		SVM	SVM + Majority vote								EMP
	Train	Test		<i>No WHEDs</i>				<i>With WHEDs</i>				
				Band50	SumBands	Sum4PCA	RCMG	Band50	SumBands	Sum4PCA	RCMG	
1—asphalt	548	6304	84.93	88.50	89.70	89.61	89.42	92.64	93.62	94.65	93.64	95.36
2—meadows	540	18146	70.79	72.73	72.51	72.83	73.33	74.45	74.49	74.50	75.09	80.33
3—gravel	392	1815	67.16	68.21	67.55	67.82	66.72	68.76	64.79	67.05	66.12	87.61
4—trees	524	2912	97.77	96.84	97.36	98.15	97.94	95.88	97.18	97.97	98.56	98.37
5—metal sheets	265	1113	99.46	99.64	99.73	99.37	99.73	99.46	99.64	98.83	99.91	99.48
6—bare soil	532	4572	92.83	94.71	94.62	95.71	94.93	97.31	97.83	98.14	97.35	63.72
7—bitumen	375	981	90.42	92.05	93.37	92.46	93.88	93.37	94.29	93.99	96.23	98.87
8—bricks	514	3364	92.78	96.82	96.49	96.11	96.20	98.42	98.78	98.42	97.92	95.41
9—shadows	231	795	98.11	96.86	96.86	96.60	98.11	95.85	95.97	96.23	96.98	97.68

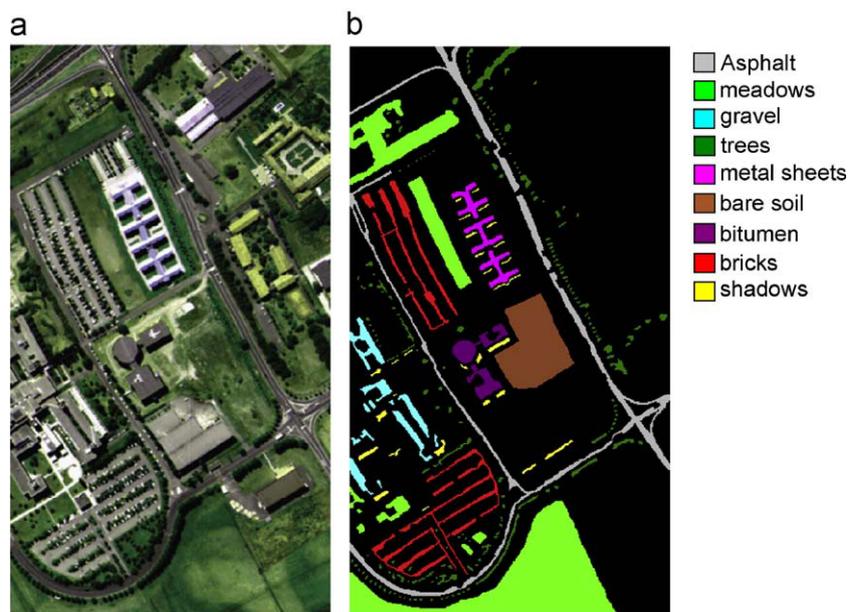


Fig. 7. *University of Pavia* image: (a) three-band color composite and (b) reference data: asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks, shadows and non-labeled pixels.

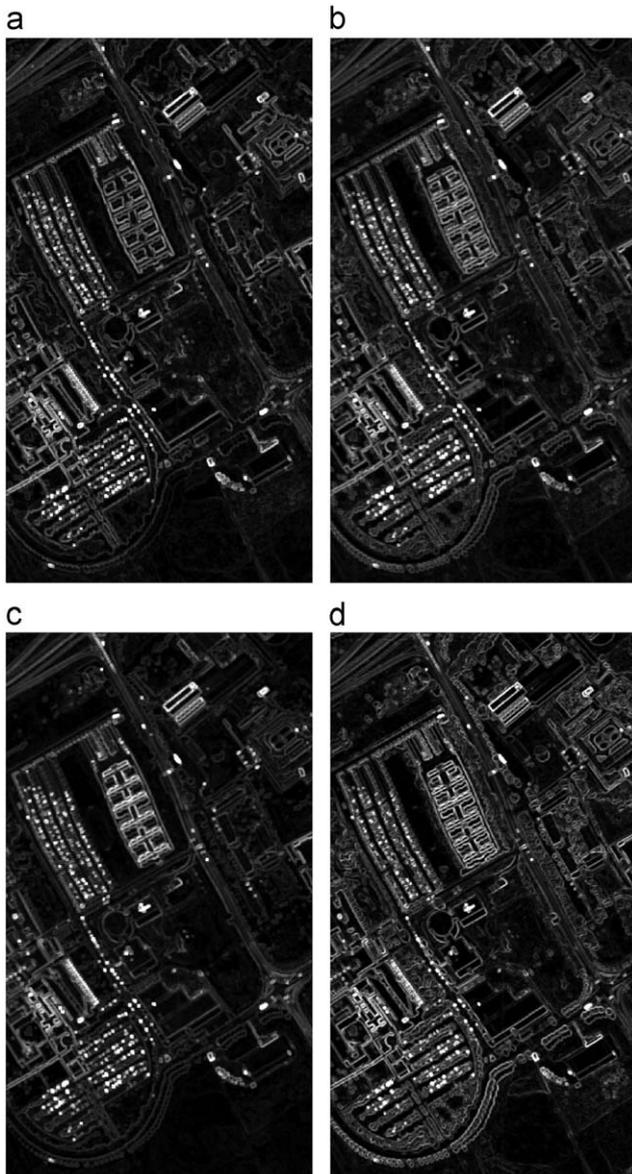


Fig. 8. Gradients of the *University of Pavia* image: (a) gradient of band no. 50; (b) sum of gradients over all bands; (c) sum of gradients for the first four PCA components; and (d) RCMG using the Euclidean distance, $r = 1$.

square structuring element E was used (center of E was in the center of the square).

2. *SumBands*: First, a morphological gradient for every band was computed. Then, the sum of gradients was obtained by (6), assuming that $\omega_\lambda = 1$, $\lambda = 1, 2, \dots, B$.
3. *Sum4PCA*: First, the PCA transformation was performed on the original image (using the ENVI software [59]). The first four principal components contained 99.16% of the total variance in the data. Morphological gradients of the first four PCA components were computed, and then summed together (using (6), with equal unitary weights).
4. *RCMG*: The RCMG on the original image was computed by (5), with $r = 1$.

The obtained four gradient images are shown in Fig. 8. The principal borders of objects are defined by all gradients. Thus, most of the spatial structures can be recognized only from the single band no. 50. However, some structures, for instance

shadows were not defined by the gradient *Band50* (see Fig. 8(a), left-bottom corner of the image). The borders of shadows are present in the other three gradient images, as in this case the information from all bands was used for computation. If we compare the images of gradients *SumBands* and *Sum4PCA* (Figs. 8(b) and (c), respectively), the latter one seems to be less noisy (for instance, see the *meadows* part in the bottom of the image). Also, the *Sum4PCA* gradient defines roads well. But as the four PCA components do not contain all the information in the original image, some details are missed in the *Sum4PCA* gradient image (for instance, *trees*). The *RCMG* gradient (see Fig. 8(d)) defines borders that are thinner and clearer than those defined by the other three gradients.

Furthermore, the watershed transformation was applied to each of the four obtained gradients, using the algorithm of Vincent and Soille [48] (based on 8-neighborhood connectivity). The resulting segmentation maps for watershed on *Band50*, *SumBands*, *Sum4PCA* and *RCMG* gradients contained 11,641, 10,558, 10,345 and 11,802 regions, respectively. Fig. 9(a) shows the segmentation map based on the *RCMG*, where the main spatial structures can be seen. (The colors of each region correspond to the label of this region, scaled in order to obtain a gray-scale 8-bit image.)

As expected, the obtained watershed results were severely oversegmented. Objects were represented mostly by several regions. The first aim was to obtain the segmentation map where each region contained pixels belonging to one object, i.e., where there were no undersegmentation errors. As mentioned in Section 2, oversegmentation can be corrected by merging regions. On the other hand, it may be more difficult to cope with the undersegmentation problem.

To assess qualitatively the accuracy of the watershed segmentation, the obtained segmentation maps were incorporated into the combined spectral-spatial classifier. First, for every segmentation map, the vector median of every region was computed, and the watershed pixels were assigned to one of the neighboring regions based on the minimal distance between the watershed pixel and the vector median of the different regions.

The multi-class pairwise (one versus one) SVM classification, with the Gaussian radial basis function (RBF) kernel, of the original hyperspectral image was performed, using the LIBSVM library [60]. The optimal parameters C and γ were determined by 5-fold cross validation: $C = 128$, $\gamma = 0.125$. Fig. 9(b) shows the obtained classification map.

After the pixel-wise SVM classification, the *majority vote* within the watershed regions was performed. Both the *No WHEDs* and *With WHEDs* approaches introduced in Section 3 were applied for each of the four segmentation maps. The global classification accuracies for the pixel-wise SVM and combined spectral-spatial classification are presented in Table 2. Here and in the following the accuracies are based on results for test data. The following measures of accuracy were used:

- *Overall accuracy* (OA) is the percentage of correctly classified pixels.
- *Average accuracy* (AA) is the mean of class-specific accuracies, i.e., the mean of the percentage of correctly classified pixels for each class.
- *Kappa coefficient* (κ) is the percentage of agreement (correctly classified pixels) corrected by the number of agreements that would be expected purely by chance [52].

The number of training and test samples along with the class-specific accuracies for test data are given in Table 1. Fig. 9(c) shows the classification map for the spectral-spatial classification, using the watershed map on the *RCMG* and the *With WHEDs* approach. In order to compare the obtained results with previous works that used

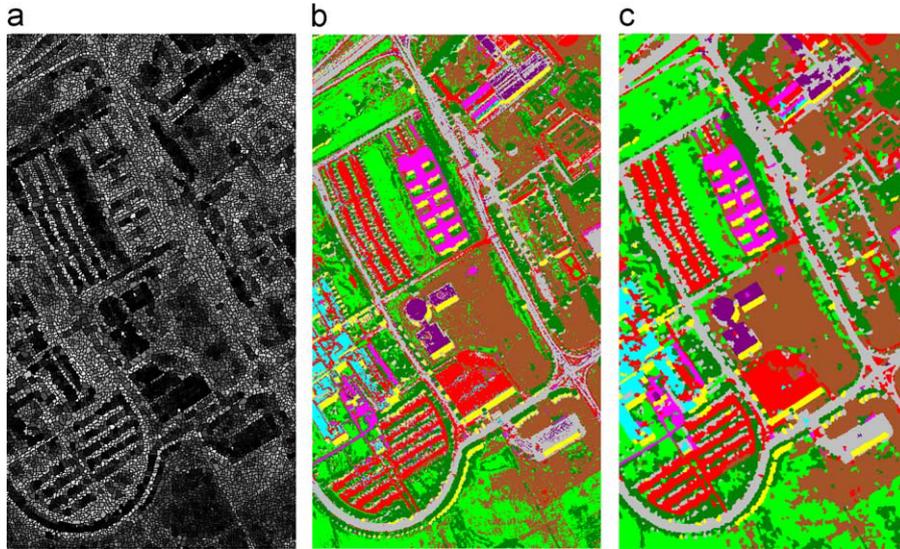


Fig. 9. *University of Pavia* image: (a) watershed on the RCMG, using the Euclidean distance, $r = 1$; (b) classification map for the SVM classification; and (c) classification map for the spectral-spatial classification (using watershed map on the RCMG, *With WHEDs* approach).

Table 2

Global classification accuracies in percentage for the *University of Pavia* image: overall accuracy (OA), average accuracy (AA) and kappa coefficient (κ).

Method		OA	AA	κ	
Pixel-wise SVM		81.01	88.25	75.86	
SVM + Majority vote	<i>No WHEDs</i>	Band50	83.00	89.59	78.31
		SumBands	83.10	89.80	78.44
		Sum4PCA	83.36	89.85	78.76
		RCMG	83.48	90.03	78.89
	<i>With WHEDs</i>	Band50	84.83	90.68	80.57
		SumBands	85.04	90.73	80.83
		Sum4PCA	85.35	91.09	81.23
		RCMG	85.42	91.31	81.30
	EMP		85.22	90.76	80.86

an SVM and spatial information for hyperspectral image classification, we have included in Tables 1 and 2 accuracies of mathematical morphology-based classification of the *University of Pavia* image using an SVM, principal components and extended morphological profiles (EMP); results are taken from Plaza et al. [61], where the same training and test samples were used for classification. This method was recently proposed by Benediktsson et al. [62] and is considered as one of the most advanced methods for spectral-spatial classification of a multi-band data. Other results of joint spectral-spatial classification of the considered image can be found in [63,15,64].

As can be seen from Table 2, all the SVM-based classifiers give high classification accuracies. Furthermore, the combining of spatial information obtained by watershed segmentation with the spectral classification results improves substantially the classification accuracies. The *With WHEDs* approach gave the best accuracies for all four segmentation results. The best global accuracies are achieved when performing the spectral-spatial classification based on the segmentation map on the RCMG and applying the *With WHEDs* approach. In that case, the overall accuracy is improved by 4.41 and the average accuracy is improved by 3.06 percentage points compared to the pixel-wise

SVM classification. The integration of the different segmentation results in the classification improves the accuracies differently: *SumBands* performs better than *Band50*, *Sum4PCA* better than *SumBands*, and RCMG performs best.

Classification accuracies were improved by the spectral-spatial classification for almost all the classes (see Table 1). For some classes, like *alphalt*, *bitumen* and *bricks*, accuracies were much improved by including spatial information, i.e., the accuracies improved by 5.81 to 9.72 percentage points, mostly because of the noise reduction. Here, different segmentation approaches led to the best classification accuracies for different classes. The RCMG segmentation map improved the classification accuracies for the classes *meadows*, *trees*, *metal sheets* and *bitumen*. For the classes *alphalt* and *bare soil*, the best accuracies were achieved when applying the *Sum4PCA* segmentation. That confirms the assumption that the *Sum4PCA* gradient defined well the information about road structures. For the classes *gravel* and *bricks*, the best accuracies were achieved using the *Band50* and *SumBands* approaches, respectively.

4.1.3. Concluding discussion

Based on the above, the following conclusions can be drawn about the accuracy of the watershed segmentation:

1. The RCMG leads to the best segmentation and classification results. Applying the vectorial gradient based on the Euclidean distance between pixel vectors to the hyperspectral image gave the best classification accuracies, despite the high-dimensionality of data.
2. The watershed segmentation based on the gradient of one band is the least accurate approach as was expected, since it may be difficult to distinguish some different neighboring structures when using only one particular band. Still, as the inclusion of the information from the *Band50* segmentation map improved the classification accuracies, most of the spatial structures could be retrieved from this single band.
3. The summing of the gradients of the first four PCA components gave slightly better results than the summing of the gradients of all bands, which indicates that the first PCA components contain the most important spatial information. Thus, applying

Table 3
Information classes, number of labeled samples, and class-specific accuracies in percentage for the *Indiana* image.

Class	No. of samples in reference data	SVM	SVM + Majority vote						HSEG
			No WHEDs			With WHEDs			
			SumBands	Sum4PCA	RCMG	SumBands	Sum4PCA	RCMG	
1—corn-no till	1434	74.59	84.04	80.94	83.89	90.24	87.84	87.45	85.82
2—corn-min till	834	64.58	78.03	78.16	74.83	90.68	83.49	88.42	84.82
3—corn	234	58.77	71.56	78.20	70.14	86.73	87.20	81.04	91.00
4—soybeans-no till	968	69.76	81.19	80.62	80.53	89.56	84.98	83.26	87.86
5—soybeans-min till	2468	79.21	89.11	90.33	89.83	99.33	97.44	98.47	98.69
6—soybeans-clean till	614	75.41	88.25	87.70	84.45	95.48	94.21	94.39	89.87
7—alfalfa	54	32.65	24.49	46.94	46.94	44.90	44.90	44.90	2.04
8—grass/pasture	497	87.05	92.19	91.52	90.63	93.53	89.96	94.42	94.42
9—grass/trees	747	92.72	97.33	94.80	96.43	99.11	98.22	99.11	98.37
10—grass/pasture-mowed	26	29.17	45.83	37.5	100	45.83	50.00	00.00	95.83
11—hay-windrowed	489	96.37	98.41	98.19	97.51	99.32	97.05	98.41	99.32
12—oats	20	22.22	5.56	16.67	5.56	00.00	00.00	00.00	00.00
13—wheat	212	90.58	95.81	95.29	94.76	99.48	99.48	99.48	98.43
14—woods	1294	91.07	93.73	92.70	94.59	97.51	95.19	97.68	98.45
15—bldg-grass-tree-drives	380	65.50	74.56	76.32	74.85	75.44	80.99	84.21	75.73
16—stone-steel towers	95	84.88	83.72	98.84	84.88	81.40	98.84	84.88	87.21

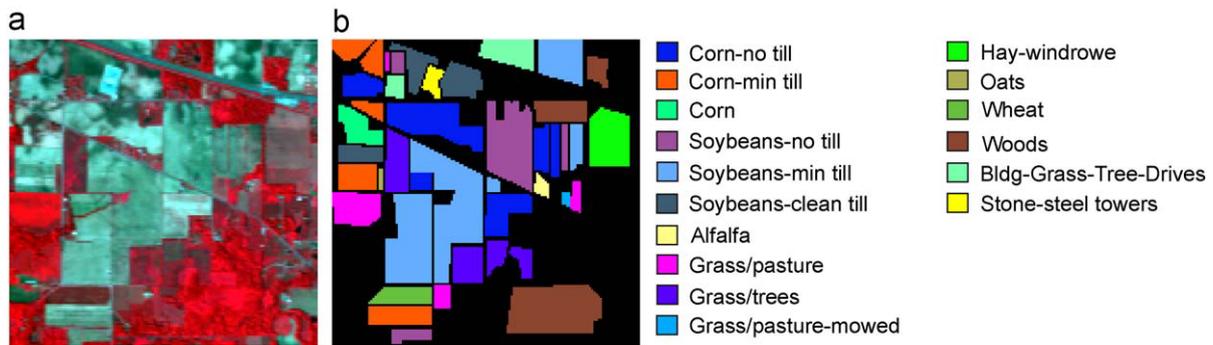


Fig. 10. *Indiana* image: (a) three-band color composite (bands 50, 27, 17); (b) reference data: corn-no till, corn-min till, corn, soybeans-no till, soybeans-min till, soybeans-clean till, alfalfa, grass/pasture, grass/trees, grass/pasture-mowed, hay-windrowed, oats, wheat, woods, bldg-grass-tree-drives, stone-steel towers and non-labeled pixels.

the feature extraction before computing the gradient and watershed leads to better segmentation results.

Thus, the incorporation of spatial information into the classifier using the *majority vote* approach led to more homogeneous objects in the resulting classification map, when compared to the pixel-wise classification. However, when performing watershed segmentation, it is usually difficult to identify small but significant structures as separate regions. They may be identified as the border pixels and then be assimilated with one of the neighboring regions. Therefore, the classification accuracies of small classes, in our case *trees* and *shadows*, are not improved significantly, or reduced.

These conclusions are confirmed by visual inspection, when comparing the classification maps of the pixel-wise versus spectral-spatial classification (see Figs. 9(b) and (c)). The spectral-spatial classification reduces significantly the noise in the classification map, resulting in more homogeneous regions in the output map. It can be also seen on the left-bottom part of the image that some small *shadows* regions were assimilated with the regions in their neighborhood.

When we compare the obtained results with the recent results of spectral-spatial classification using SVM and EMPs (see Tables 1

and 2), the proposed segmentation and classification approach leads to higher global accuracies. Furthermore, accuracies for 4 from 9 classes are improved by our technique.

4.2. Segmentation and classification of the *Indiana* image

4.2.1. Dataset

In the second case study, the developed segmentation and classification algorithms are tested on a hyperspectral image of a rural area (the *Indiana* image) with more bands and a lower spatial resolution as compared to the *University of Pavia* image. The *Indiana* image was captured by the AVIRIS sensor over the Indian Pine test site in Northwestern Indiana [45]. The image is 145 by 145 pixels, and the spatial resolution is 20 m per pixel. It is composed of 220 spectral channels, and the full spectral range was used for experiments. The data contains 16 classes, which are detailed in Table 3, with a number of samples for each class in the available reference data. Fig. 10 depicts the 3-band false color composite and the reference data. In our experiments, we have chosen randomly 10% of the samples for each class from the reference data as training samples (in order to test classification performances when limited reference data are available).

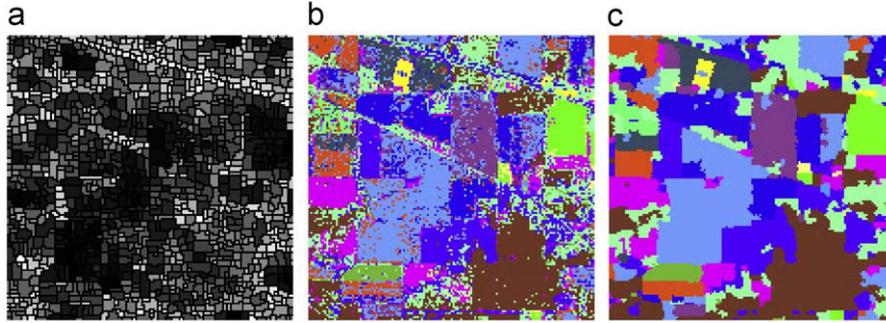


Fig. 11. *Indiana* image: (a) watershed on the sum of gradients over all bands; (b) classification map for the SVM classification; and (c) classification map for the spectral-spatial classification (using *SumBands* method, *With WHEDs* approach).

Table 4

Global classification accuracies in percentage for the *Indiana* image: Overall Accuracy (OA), Average Accuracy (AA) and kappa coefficient (κ).

Method		OA	AA	κ	
Pixel-wise SVM		78.76	69.66	75.73	
SVM + Majority Vote	<i>No WHEDs</i>	<i>SumBands</i>	87.03	75.24	85.18
		<i>Sum4PCA</i>	86.92	77.79	85.05
		<i>RCMG</i>	86.79	79.36	84.88
	<i>With WHEDs</i>	<i>SumBands</i>	93.78	80.53	92.88
		<i>Sum4PCA</i>	91.67	80.61	90.47
		<i>RCMG</i>	92.48	77.26	91.39
HSEG		92.20	80.49	91.06	

4.2.2. Experimental results

In this experiment, we do not choose one single band to perform a gradient, because there are more classes, with similar spectral responses, and some classes can be confused using only one specific band. Thus, three gradient functions were computed: *SumBands*, *Sum4PCA*, *RCMG*. They were computed in the same way as for the *University of Pavia* image, as described in the previous subsection. For the *Indiana* image, the first four PCA components contain 99.23% of the total variance for the data.

As for the previous data set, the watershed transformation was applied to each of the gradient functions. The segmentation maps contain 1215, 1097 and 1277 regions for the *SumBands*, *Sum4PCA* and *RCMG* gradients, respectively. Fig. 11(a) shows the segmentation map obtained from the *SumBands* gradient. It is difficult to evaluate the accuracy of segmentation from the image, since it is strongly oversegmented. In order to perform such an evaluation, the information from the segmentation maps was used for the spectral-spatial classification, as described below.

First, the segmentation maps for the *With WHEDs* approach were computed, by assigning the watershed pixels to the neighboring regions that had the closest vector median. Then, the multi-class one versus one SVM classification, with the Gaussian RBF kernel, of the original *Indiana* image was performed. The parameters C and γ were computed by 5-fold cross validation giving: $C = 1024$, $\gamma = 2^{-7}$. The resulting classification map is given in Fig. 11(b). The classification accuracies for test samples are presented in Tables 3 and 4.

The low spatial resolution of the *Indiana* image leads to the presence of highly mixed pixels which complicates the classification problem. Furthermore, some classes represent small crop fields, and the number of samples in the reference data for different classes varies from 20 to 2468 samples per class. To train the SVM classifier, 10% of the samples for each class were chosen

randomly from the reference data. Therefore, some classes were represented by only a few samples in the training set (2 samples for the class *oats*), which may not provide sufficient representatives for the small classes.

Despite these problems, 78.76% of the pixels in the test set were correctly classified by the SVM classifier. As expected, the classification accuracies for the classes that were represented by only a few training samples are low (for instance, only 22.22% accuracy was observed for the class *oats*). The low classification accuracies for the classes *alfalfa*, *grass/pasture-mowed* and *oats* reduced the average classification accuracy to 69.66%.

Now we turn to the spectral-spatial classification. As in the first experiment, the watershed segmentation maps were used in the spectral-spatial classification based on the *majority vote* method, with the *No WHEDs* and *With WHEDs* approaches.

Tables 3 and 4 give class-specific and global classification accuracies for the spectral-spatial classification of the *Indiana* image, respectively. Fig. 11(c) shows the classification map for the spectral-spatial classification, using the watershed map on the *SumBands* gradient and the *With WHEDs* approach. Previous classification results for the *Indiana* image can be found in [65,61] for comparison. However, the accuracies in the referenced works are not directly compared with those given in this paper, because different training-test sets are used. Furthermore, in order to evaluate the efficiency of the proposed segmentation and classification scheme using watershed, we have included in Tables 3 and 4 the spectral-spatial classification results, obtained by applying another segmentation technique followed by the procedure proposed in this paper (SVM classification and majority vote within the regions from a segmentation map). This time, the hierarchical image segmentation (HSEG) algorithm [24] was used to segment a hyperspectral image. The NASA Goddard's RHSEG software provides an efficient implementation of this algorithm [66], which we used for our investigation. At the initialization step each pixel is considered as one region. The algorithm merges iteratively the most similar adjacent and non-adjacent regions. We used a Spectral Angle Mapper between the region mean vectors as the dissimilarity criterion between regions. The relative importance of merging of non-adjacent regions *versus* region growing (when only adjacent regions are merged) can be tuned. From experimental results, simple region growing leads to the highest classification accuracies. Segmentation maps for several levels of hierarchy were chosen interactively, and the results leading to the best classification are shown.

As in the previous experiment, the combined spectral-spatial classification improved the classification accuracies, and the *With WHEDs* approach gave the best performances. The use of the watershed map based on the *SumBands* gradient led to the highest overall accuracy (an improvement by 15.02 percentage points compared to the pixel-wise classification) and kappa coefficient, while the highest average accuracy was achieved by using the

Sum4PCA gradient (it is 10.95 percentage points higher than for the pixel-wise classification). The watershed based on the *RCMG* gradient led to lower classification accuracies than the one based on the *SumBands* gradient. And the *RCMG* approach gave a higher overall accuracy but a lower average accuracy than the *Sum4PCA* approach.

The class-specific accuracies were improved by using the spatial information in the classification for almost all the classes (see Table 3). After the combined spectral-spatial classification, the classification accuracy was reduced for only one class (*oats*). The classification using the *With WHEDs* approach led to 0% accuracy for this class. The *oats* class has only 20 pixels in the reference map, and represents a small and very narrow (of 2 pixels width) rectangular field. Thus, after majority voting within the watershed regions, pixels of the *oats* class were assigned to the classes in its neighborhood (mostly to the *grass/trees* class). This drawback of spectral-spatial classification, i.e., the risk of losing small spatial structures, was discussed in the previous subsection.

4.2.3. Concluding discussion

In assessing the watershed segmentation approaches we note that the results for the *Indiana* image are different from those for the *University of Pavia* image. For the *Indiana* image, the linear combination (sum) of gradients of all bands or several principal components leads to better segmentation results than the use of the vectorial gradient. This can be explained by the fact that the *Indiana* image contains more spectral channels (220 channels versus 103 channels for the *University of Pavia* image). Furthermore, the vectorial gradient, based on the Euclidean distance between pixel vectors does not give accurate results due to the curse of dimensionality. In this case, it appears to be more appropriate to compute marginal gradients of every band and to sum them together or to reduce the spectral dimension by performing feature extraction first, and then compute a gradient.

The spectral-spatial classification improved classification accuracies when compared to pixel-wise classification. The improvement of the global accuracies was more significant than for the *University of Pavia* image. From a visual observation (see Figs. 11(b) and (c)), the classification map obtained by the spectral-spatial classification is seen to be much less noisy than the one obtained by the pixel-wise classification. From the obtained results, we can draw the conclusion that the *Indiana* image contains a lot of relatively large regularly structured spatial objects (crop fields) which is the reason why the inclusion of the spatial information was very efficient and improved the classification accuracies significantly.

If we analyze classification results obtained by majority vote within the HSEG regions, we can conclude that:

1. The global and most of class-specific accuracies are improved when compared to pixel-wise classification. This proves the efficiency of the proposed spectral-spatial classification scheme.
2. The accuracies are lower than those obtained when using watershed for segmentation (except for 2 classes). This confirms why it is desirable to use watershed segmentation for hyperspectral images.

5. Conclusions

Hyperspectral imaging provides rich spectral information per pixel, increasing the capability to distinguish physical structures in a scene. However, a large number of spectral channels presents

challenges to image classification. While pixel-wise classification techniques process each pixel independently without considering information about spatial structures, further improvement of classification performances can be achieved by the incorporation of spatial information into classifier, especially in areas where structural information is important to distinguish between classes.

In this paper, there are two main contributions:

1. The extension of the watershed segmentation algorithm for hyperspectral images was proposed. In particular, different ways were investigated to obtain a one-band gradient function from a hyperspectral image and the combination of watershed segmentation maps was proposed.
2. A new methodology, a spectral-spatial classification scheme for hyperspectral images was proposed. The new method is based on the pixel-wise SVM classification, followed by majority voting within the watershed regions. Thus, segmentation defines an adaptive neighborhood for each pixel.

In two experiments it was shown that the combined spectral-spatial classification, based on majority voting within the regions obtained by the watershed segmentation algorithms, led to higher classification accuracies when compared to pixel-wise classification or previously proposed techniques. Furthermore, classification maps with more homogeneous regions were obtained with the proposed approach.

In conclusion, the proposed classification methodology succeeded in taking advantage of spatial and spectral information simultaneously. Furthermore, experimental results have revealed that it is desirable to use the watershed segmentation map in a spectral-spatial classifier. The watershed transformation is a completely unsupervised method since it does not require any input parameters (thresholds). Therefore, it can be incorporated into an automatic classification system.

The proposed spectral-spatial classification scheme is especially suitable in classifying images with large spatial structures. The drawback of the proposed method is that when applying watershed segmentation, small spatial structures are often not identified as separate regions. This leads to the assimilation of these structures with larger neighboring structures when majority voting is performed within the watershed regions.

In our future work, we will attempt to improve the segmentation results. In particular, we envision to explore the use of additional filtering and merging of regions for that purpose.

Acknowledgments

This research is supported in part by the Marie Curie Research Training Network "HYPER-I-NET". The authors would like to thank Paolo Gamba and David Landgrebe for providing the hyperspectral data and James Tilton for providing the NASA Goddard's RHSEG software.

References

- [1] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic Publishers, Dordrecht, 2003.
- [2] R.O. Green, M.L. Eastwood, C.M. Sarture, T.G. Chrien, M. Aronsson, B.J. Chippendale, J.A. Faust, B.E. Pavri, C.J. Chovit, M.S. Solis, M.R. Olah, O. Williams, *Imaging spectroscopy and the airborne visible infrared imaging spectrometer (AVIRIS)*, *Remote Sens. Environ.* 65 (3) (1998) 227–248.

- [3] G. Hughes, On the mean accuracy of statistical pattern recognizers, *IEEE Trans. Inf. Theory* 14 (1) (1968) 55–63.
- [4] D. Landgrebe, Hyperspectral image data analysis, *IEEE Signal Process. Mag.* 1053–5888 (2002) 17–28.
- [5] D.A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, Wiley, New York, 2003.
- [6] C.H. Chen, P.-G. Peter Ho, Statistical pattern recognition in remote sensing, *Pattern Recognition* 41 (9) (2008) 2731–2741.
- [7] N. Bali, A. Mohammad-Djafari, Bayesian approach with hidden Markov modeling and mean field approximation for hyperspectral data analysis, *IEEE Trans. Image Process.* 17 (2) (2008) 217–225.
- [8] S. Subramanian, N. Gat, M. Sheffield, J. Barhen, N. Toomarian, Methodology for hyperspectral image classification using novel neural network, in: *Proceedings of SPIE*, vol. 3071, 1997, pp. 128–137.
- [9] H. Yang, F.V.D. Meer, W. Bakker, Z.J. Tan, A back-propagation neural network for mineralogical mapping from AVIRIS data, *Int. J. Remote Sens.* 20 (1) (1999) 97–110.
- [10] C. Hernández-Espinosa, M. Fernández-Redondo, J. Torres-Sospedra, Some experiments with ensembles of neural networks for classification of hyperspectral images, in: *ISNN* (1), 2004, pp. 912–917.
- [11] P.K. Goel, S.O. Prasher, R.M. Patel, J.A. Landry, R.B. Bonnell, A.A. Viau, Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn, *Comput. Electron. Agri.* 39 (2003) 67–93.
- [12] H. Zhou, Z. Mao, D. Wang, Classification of coastal areas by airborne hyperspectral image, in: *Proceedings of SPIE*, vol. 5832, 2005, pp. 471–476.
- [13] C. Vaiphasa, Innovative genetic algorithm for hyperspectral image classification, in: *Proceedings of the International Conference Map Asia*, vol. 20, 2003.
- [14] G. Camps-Valls, L. Bruzzone, Kernel-based methods for hyperspectral image classification, *IEEE Trans. Geosci. Remote Sens.* 43 (6) (2005) 1351–1362.
- [15] M. Fauvel, Spectral and spatial methods for the classification of urban remote sensing data, Ph.D. Thesis, Grenoble Institute of Technology, 2007.
- [16] M. Fauvel, J. Chanussot, J.A. Benediktsson, Evaluation of kernels for multiclass classification of hyperspectral remote sensing data, in: *Proceedings of ICASSP2006*, vol. 2, 2006, pp. II-813–II-816.
- [17] B. Guo, S.R. Gunn, R.I. Dampier, J.D.B. Nelson, Customizing kernel functions for SVM-based hyperspectral image classification, *IEEE Trans. Image Process.* 17 (4) (2008) 622–629.
- [18] M. Pesaresi, J.A. Benediktsson, A new approach for the morphological segmentation of high-resolution satellite imagery, *IEEE Trans. Geosci. Remote Sens.* 39 (2) (2001) 309–320.
- [19] A. Farag, R. Mohamed, A. El-Baz, A unified framework for map estimation in remote sensing image segmentation, *IEEE Trans. Geosci. Remote Sens.* 43 (7) (2005) 1617–1634.
- [20] K. Fu, J. Mui, A survey on image segmentation, *Pattern Recognition* 13 (1) (1981) 3–16.
- [21] Y. Tarabalka, J. Chanussot, J.A. Benediktsson, J. Angulo, M. Fauvel, Segmentation and classification of hyperspectral data using watershed, in: *Proceedings of IGARSS'08*, Boston, USA, 2008, pp. III-652–III-655.
- [22] Y. Tarabalka, J.A. Benediktsson, J. Chanussot, J. Angulo, M. Fauvel, Classification of hyperspectral data using Support Vector Machines and adaptive neighborhoods, in: *Proceedings of the Sixth EARSeL SIG IS workshop*, Tel Aviv, Israel, 2009.
- [23] A. Darwish, K. Leukert, W. Reinhardt, Image segmentation for the purpose of object-based classification, *Proceedings of IGARSS'03*, vol. 3, 2003, pp. 2039–2041.
- [24] J. Tilton, Analysis of hierarchically related image segmentations, *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data* (2003) 60–69.
- [25] P. Soille, Morphological partitioning of multispectral images, *J. Electron. Imaging* 5 (3) (1996) 252–265.
- [26] G. Flouzat, O. Amram, S. Cherchali, Spatial and spectral segmentation of satellite remote sensing imagery using processing graphs by mathematical morphology, in: *Proceedings of IGARSS '98*, vol. 4, 1998, pp. 1769–1771.
- [27] P. Scheunders, Multivalued image segmentation based on first fundamental form, in: *Proceedings of the 11th International Conference on Image Analysis and Processing*, 2001, pp. 185–190.
- [28] P. Li, X. Xiao, Evaluation of multiscale morphological segmentation of multispectral imagery for land cover classification, *Proceedings of IGARSS'04*, vol. 4, 2004, pp. 2676–2679.
- [29] G. Noyel, J. Angulo, D. Jeulin, Morphological segmentation of hyperspectral images, *Image Anal. Stereol.* 26 (2007) 101–109.
- [30] H.G. Akçay, S. Aksoy, Automatic detection of geospatial objects using multiple hierarchical segmentations, *IEEE Trans. Geosci. Remote Sens.* 46 (7) (2008) 2097–2111.
- [31] J. Chanussot, P. Lambert, Bit mixing paradigm for multivalued morphological filters, in: *Proceedings of IEE IPA'97*, 1997, pp. 804–808.
- [32] J. Chanussot, P. Lambert, Total ordering based on space filling curves for multivalued morphology, in: *Proceedings of ISMM'98*, 1998, pp. 51–58.
- [33] P. Lambert, J. Chanussot, Extending mathematical morphology to color image processing, in: *Proceedings of CGIP'00*, 2000, pp. 158–163.
- [34] A.G. Hanbury, J. Serra, Morphological operators on the unit circle, *IEEE Trans. Image Process.* 10 (12) (2001) 1842–1850.
- [35] J. Angulo, J. Serra, Morphological coding of color images by vector connected filters, in: *Proceedings of ISSPA'2003*, vol. 1, 2003, pp. 69–72.
- [36] E. Aptoula, S. Lefèvre, A comparative study on multivariate mathematical morphology, *Pattern Recognition* 40 (11) (2007) 2914–2929.
- [37] L. Shafarenko, M. Petrou, J.V. Kittler, Automatic watershed segmentation of randomly textured color images, *IEEE Trans. Image Process.* 6 (11) (1997) 1530–1544.
- [38] J. Chanussot, P. Lambert, Watershed approaches for color image segmentation, in: *Proceedings of IEEE NSIP'99*, 1999, pp. 129–133.
- [39] J. Angulo, J. Serra, Mathematical morphology in color spaces applied to the analysis of cartographic images, in: *Proceedings of the GEOPRO 2003—International Workshop Semantic Processing of Spatial Data*, 2003, pp. 59–66.
- [40] J. Angulo, Unified morphological color processing framework in a lum/sat/hue representation, in: *Proceedings of ISMM'2005*, 2005, pp. 387–396.
- [41] S.V. der Linden, A. Janz, B. Waske, M. Eiden, P. Hostert, Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines, *J. Appl. Remote Sens.* 1 (1) (2007) 013543.
- [42] A. Widayati, B. Verbist, A. Meijerink, Application of combined pixel-based and spatial-based approaches for improved mixed vegetation classification using IKONOS, in: *Proceedings of the 23th Asian Conference on Remote Sensing*, 2002, 8pp.
- [43] R. Cook, I. McConnell, C.J. Oliver, MUM (merge using moments) segmentation for SAR images, in: *Proceedings of SPIE*, vol. 2316, 1994, pp. 92–103.
- [44] P. Soille, *Morphological Image Analysis*, second ed., Springer, Berlin, 2003.
- [45] AVIRIS NW Indiana's Indian Pines 1992 data set [Online], Available: <ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/92AV3C> (original files) and <ftp://ftp.ecn.purdue.edu/biehl/PCMultiSpec/ThyFiles.zip> (ground truth).
- [46] F. Meyer, S. Beucher, Morphological segmentation, *J. Vis. Commun. Image Represent.* 1 (1990) 21–46.
- [47] S. Beucher, C. Lantuejoul, Use of watersheds in contour detection, in: *International Workshop Image Processing, Real-Time Edge and Motion Detection/ Estimation*, Rennes, France, CCETT/IRISA, 1979.
- [48] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (6) (1991) 583–598.
- [49] F. Meyer, Topographic distance and watershed lines, *Signal Process.* 38 (1) (1994) 113–125.
- [50] S. Stoev, Rafsi—a fast watershed algorithm based on rainfalling simulation, in: *Proceedings of the Eighth International Conference on Computer Graphics, Visualization, and Interactive Digital Media*, 2000.
- [51] V. Osma-Ruiz, J.I. Godino-Llorente, N. Sáenz-Lechón, P. Gómez-Vilda, An improved watershed algorithm based on efficient computation of shortest paths, *Pattern Recognition* 40 (3) (2007) 1078–1090.
- [52] J.A. Richards, X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, New York, 1999.
- [53] M.R. Gupta, N.P. Jacobson, Wavelet principal component analysis and its application to hyperspectral images, in: *Proceedings of IEEE ICIP'06*, 2006, pp. 1585–1588.
- [54] A.A. Green, M. Berman, P. Switzer, M.D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Trans. Geosci. Remote Sens.* 26 (1) (1988) 65–74.
- [55] J. Wang, C.-I. Chang, Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis, *IEEE Trans. Geosci. Remote Sens.* 44 (6) (2006) 1586–1600.
- [56] S. Moussaoui, H. Hauksdottir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Dout'e, J.A. Benediktsson, On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation, *Neurocomputing* 71 (2008) 2194–2208.
- [57] A.N. Evans, X.U. Liu, A morphological gradient approach to color edge detection, *IEEE Trans. Image Process.* 15 (6) (2006) 1454–1463.
- [58] J. Astola, P. Haavisto, Y. Neuvo, Vector median filters, *Proc. IEEE* 78 (4) (1990) 678–689.
- [59] ITT Corporation, For further information about the ENVI software: <http://rsinc.com/envi/>.
- [60] C. Chang, C. Lin, LIBSVM—a library for support vector machines, Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (2008).
- [61] A. Plaza, J.A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J.A. Gualtieri, M. Marconcini, J.C. Tilton, G. Trianni, Recent advances in techniques for hyperspectral image processing, *Remote Sens. Environ.* 113 (Suppl. 1) (2009) S110–S122.
- [62] J.A. Benediktsson, J.A. Palmason, J.R. Sveinsson, Classification of hyperspectral data from urban areas based on extended morphological profiles, *IEEE Trans. Geosci. Remote Sens.* 43 (3) (2005) 480–491.
- [63] S. Aksoy, Spatial techniques for image classification, in: C.H. Chen (Ed.), *Signal and Image Processing for Remote Sensing*, Taylor Francis, 2006, pp. 491–513.
- [64] M. Fauvel, J. Chanussot, J.A. Benediktsson, J.R. Sveinsson, Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles, *IEEE Trans. Geosci. Remote Sens.* 46 (10) (2008).
- [65] S. Tadjudin, D. Landgrebe, Classification of high dimensional data with limited training samples, Ph.D. Thesis, School of Electrical and Computer Engineering, Purdue University, 1998.
- [66] J.C. Tilton, HSEG/RHSEG, HSEGViewer and HSEGReader user's manual (version 1.40), Provided with the evaluation version of RHSEG available from: <http://ipp.gsfc.nasa.gov/RHSEG>.

About the Author—YULIYA TARABALKA received her Bachelor degree in Computer Science (2005) from the Ternopil Ivan Pul'uj State Technical University, Ukraine and M.Sc. degree in Signal and Image Processing (2007) from the Grenoble Institute of Technology (INPG), France. From July 2007 to January 2008 she worked as a Researcher with the Norwegian Defence Research Establishment, Norway. Since 2007, she is a Ph.D. student, pursuing a co-joint degree between University of Iceland, Iceland and INPG, France. Her current research work is funded by the "HYPER-I-NET" Marie Curie Research Training Network. Her research interests are in the areas of image processing, pattern recognition, hyperspectral imaging and development of efficient algorithms.

About the Author—JOCELYN CHANUSSOT received the M.Sc. degree in Electrical Engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA-French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab, INPG). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. Dr. Chanussot is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing. He was an Associate Editor for the IEEE Geoscience and Remote Sensing Letters (2005–2007) and for Pattern Recognition (2006–2008). He is the Cochair of the GRS Data Fusion Technical Committee (2005–2008) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008). He is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007) and a member of the IEEE Geoscience and Remote Sensing AdCom (2009–2011). He is the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS).

About the Author—JÓN ATLI BENEDIKTSSON received the Cand.Sci. degree in Electrical Engineering from the University of Iceland, Reykjavik, in 1984, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively. He joined the Department of Electrical and Computer Engineering, University of Iceland in 1991 and is currently Pro Rector for Academic Affairs and Professor of Electrical and Computer Engineering at the University of Iceland. He has held visiting positions at the Department of Information and Communication Technology, University of Trento, Trento, Italy (2002–present), School of Computing and Information Systems, Kingston University, Kingston upon Thames, U.K. (1999–2004), the Joint Research Centre of the European Commission, Ispra, Italy (1998), Denmark's Technical University (DTU), Lyngby (1998), and the School of Electrical and Computer Engineering, Purdue University (1995). He was a Fellow at the Australian Defence Force Academy, Canberra, A.C.T., Australia, in August 1997. From 1999 to 2004, he was a Chairman of the energy company Metan Ltd. His research interests are in remote sensing, pattern recognition, neural networks, image processing, and signal processing, and he has published extensively in those fields. Dr. Benediktsson is Editor of the IEEE Transactions on Geoscience and Remote Sensing (TGARS) and Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He was Associate Editor of TGARS from 1999 to 2002. He co-edited (with Prof. D. A. Landgrebe) a Special Issue on Data Fusion of TGARS (May 1999) and (with P. Gamba and G. G. Wilkinson) a Special Issue on Urban Remote Sensing from Satellite (October 2003). From 1996 to 1999, he was the Chairman of GRSS Technical Committee on Data Fusion and was elected to the Administrative Committee of the GRSS in 2000. He was GRSS Vice President of Technical Activities in 2002, GRSS Vice President of Professional Activities in 2008–2009 and elected GRSS Executive Vice President for 2010–2011. He was the founding Chairman of the IEEE Iceland Section and served as its Chairman from 2000 to 2003. Currently, he is the Chairman of the University of Iceland's Quality Assurance Committee (2006–). He was the Chairman of the University of Iceland's Science and Research Committee (1999–2005), a member of Iceland's Science and Technology Council (2003–2006), and a member of the Nordic Research Policy Council (2004). He was a member of a NATO Advisory Panel of the Physical and Engineering Science and Technology SubProgramme (2002–2003). Dr. Benediktsson is a Fellow of the IEEE. He received the Stevan J. Kristof Award from Purdue University in 1991 as outstanding graduate student in remote sensing. In 1997, he was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award, in 2000, he was granted the IEEE Third Millennium Medal, in 2004, he was a co-recipient of the University of Iceland's Technology Innovation Award, in 2006 he received the yearly research award from the Engineering Research Institute of the University of Iceland, and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a co-founder of the start up company Oxymap and member of Societas Scientiarum Islandica and Tau Beta Pi.

Paper 3

Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2973-2987, Aug. 2009.

Spectral–Spatial Classification of Hyperspectral Imagery Based on Partitional Clustering Techniques

Yuliya Tarabalka, *Student Member, IEEE*, Jón Atli Benediktsson, *Fellow, IEEE*, and Jocelyn Chanussot, *Senior Member, IEEE*

Abstract—A new spectral–spatial classification scheme for hyperspectral images is proposed. The method combines the results of a pixel wise support vector machine classification and the segmentation map obtained by partitional clustering using majority voting. The ISODATA algorithm and Gaussian mixture resolving techniques are used for image clustering. Experimental results are presented for two hyperspectral airborne images. The developed classification scheme improves the classification accuracies and provides classification maps with more homogeneous regions, when compared to pixel wise classification. The proposed method performs particularly well for classification of images with large spatial structures and when different classes have dissimilar spectral responses and a comparable number of pixels.

Index Terms—Clustering, hyperspectral images, majority vote, segmentation, spectral–spatial classification.

I. INTRODUCTION

THE ACCURATE classification of remote sensing images is an important task for many practical applications, such as precision agriculture, monitoring and management of the environment, and security and defense issues. The advent and growing availability of hyperspectral imagery, which records hundreds of spectral bands, has opened new possibilities in image analysis and classification. Examples of hyperspectral imaging systems are Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [1], HYDICE [2], ARCHER [3], HyMap [4], and Hyperion [5]. They cover a range of 126–512 spectral channels, with the spatial resolution of 3–30 m per pixel. Thus, every pixel in a hyperspectral image contains values that correspond to the detailed spectrum of reflected light [6]. This rich spectral information in every spatial location increases the capability to distinguish different physical materials and objects, leading to the potential of a more accurate image classification.

Manuscript received October 3, 2008; revised December 16, 2008. First published April 24, 2009; current version published July 23, 2009. This work was supported in part by the Marie Curie Research Training Network “HYPER-I-NET.”

Y. Tarabalka is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland and also with the GIPSA-Lab-Grenoble Institute of Technology, Domaine Universitaire, 38402 Saint-Martin-d’Hères Cedex, France (e-mail: yuliya.tarabalka@hyperinet.eu).

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

J. Chanussot is with GIPSA-Lab-Grenoble Institute of Technology, Domaine Universitaire, 38402 Saint-Martin-d’Hères Cedex, France (e-mail: jocelyn.chanussot@lis.inpg.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2009.2016214

An extensive literature is available on the classification of hyperspectral images where a wide range of pixel-level processing techniques is proposed, i.e., techniques that assign each pixel to one of the classes based on its spectral values. Maximum-likelihood or Bayesian estimation methods [7], decision trees [8], [9], neural networks [10]–[12], genetic algorithms [13], and kernel-based techniques [14], [15] have been investigated for this purpose. In particular, support vector machines (SVMs) have shown a good performance for classifying high-dimensional data when a limited number of training samples are available [14], [16], [17].

To improve classification results, the contextual information should be considered for incorporation into the classifiers. Spectral–spatial classification aims at assigning each image pixel to one class using a feature vector based on the following: 1) its own spectral value (the spectral information) and 2) information extracted from its neighborhood (referred to as the spatial information in the following). One of the approaches of spectral–spatial classification consists in including the information from the closest neighborhood to classify each pixel. These fixed-window-based methods that use morphological filtering [15], morphological leveling [18], [19], or Markov random fields [20] have shown improvements in classification accuracies compared to the pixel wise methods, when applied to hyperspectral images. However, the use of these methods raises the problem of scale selection, particularly when small or complex structures are present in the image.

Another approach to include spatial information in classification consists in performing image segmentation. *Segmentation* can be defined as an exhaustive partitioning of the input image into regions, each of which is considered to be homogeneous with respect to some criterion of interest (homogeneity criterion, e.g., intensity or texture) [21]. These regions form a segmentation map that can be used as spatial structures for a spectral–spatial classification.

In this paper, we propose a new *spectral–spatial classification scheme* for hyperspectral data. The proposed method combines the results of a pixel wise spectral classification and a segmentation map, aiming to improve classification accuracies, when compared to pixel wise classification only.

Fu and Mui [22] identified three classes of image segmentation techniques: *edge-based*, *region-based*, and *characteristic feature thresholding* or *clustering*. Lambert and Macaire [23] have split the last class into two: *histogram-based* and *cluster-based* methods. *Edge-based* techniques search for discontinuities in the image, while *region-based* techniques search for similarities between image regions. Methods from these two

classes operate in the spatial domain. Their adaptation to the multidimensional images is a challenging task.

The other two, *histogram-based* and *cluster-based* techniques, work in the spectral domain. They search for similarities between image pixels and clusters of pixels, not taking into consideration the spatial location of these pixels. The *histogram-based* methods relate the modes of a spectral histogram to homogeneous regions in the image [24], [25]. With a high dimensionality, these methods become memory consuming and produce less accurate results. The *cluster-based* segmentation techniques aim at finding distinct structures in the spectral feature space. Thus, *clustering* is an exhaustive partitioning of a set of pixels from the input image into homogeneous groups of pixels. In this paper, the *cluster-based* segmentation of hyperspectral images will be explored.

A taxonomy and survey of clustering techniques can be found in [26]. Two principal groups of clustering methods can be distinguished: hierarchical and partitional approaches. Hierarchical methods usually produce a dendrogram, where at the lowest level, each cluster contains only one pixel (i.e., each pixel forms a cluster), and with the increase of levels, the most similar clusters are merged (and the number of clusters decreases). Then, the result with the desired number of clusters can be chosen. Lee and Crawford [27] have applied the hierarchical clustering approach for unsupervised classification of hyperspectral images. Although hierarchical clustering is a versatile technique for image segmentation that can produce a series of segmentation results, it requires a lot of computational time. Its application to high-dimensional data leads to significant time and memory requirements, and it becomes more difficult to cope with large dendrograms.

In this paper, the use of partitional clustering for hyperspectral image segmentation is investigated. Two algorithms are considered for this purpose: ISODATA (squared-error clustering method) [28] and expectation maximization (EM) for the Gaussian mixture resolving [29]. These algorithms produce a single partition of the data, and the number of desired clusters must be chosen. The other problem accompanying the use of these techniques is that the clustering results depend on the initialization (it will be discussed in the next sections). However, the computational complexity of these algorithms is lower than that in the case of hierarchical clustering. Moreover, efficient implementations are possible. Venkateswarlu and Raju [30] proposed an algorithm to speed up the ISODATA algorithm. Tarabalka *et al.* [31] have shown that the parallel implementation of the EM algorithm on the graphical processing unit is feasible and efficient.

The results of hyperspectral image segmentation are further incorporated into a spectral–spatial classifier. The SVM classifier is used in the proposed method. In previous studies, the integration of spectral and spatial information into classifier with the use of SVM was achieved in different ways: Within the framework of composite kernels, spectral and contextual information was combined using composite kernels and then each pixel was classified [19], [32]–[35]. Van der Linden *et al.* [36] used another approach which consisted of computing a vector mean for each region (such that the value in each spectral channel represented the average spectral information of the

pixels in this region in the respective channel) and then using this vector as a feature vector to classify each region by an SVM classifier. The use of composite kernels led to the improvement of the classification accuracies when compared to an SVM classification using spectral information only; however, the approach of classifying regions using their vector means did not show any improvement over results obtained by using only a spectral-based pixel wise SVM classification.

Here, we propose a new spectral–spatial classification scheme, where pixel wise SVM classification and segmentation by clustering are performed independently, and then, the results are combined using the *majority vote* approach [37]. Thus, the segmentation defines an *adaptive neighborhood* for each pixel. These neighborhoods are then used for the contextual regularization following a spectral pixel wise classification. Finally, a spatial postregularization (PR) of the classification map is performed.

Although the proposed scheme has been designed for hyperspectral images, the method is general and can be applied for other types of data as well.

Two hyperspectral airborne images were used to demonstrate experimental results: a 103-band Reflective Optics System Imaging Spectrometer (ROSIS) image of the University of Pavia, Italy, and a 220-band AVIRIS image taken over the Northwestern Indiana’s Indian Pine site [38].

The outline of this paper is as follows. In Section II, segmentation of hyperspectral data using partitional clustering techniques is discussed. Section III describes the proposed spectral–spatial classification scheme. Experimental results are discussed in Section IV. Finally, conclusions are drawn in Section V.

II. SEGMENTATION OF HYPERSPECTRAL DATA BY PARTITIONAL CLUSTERING

In this section, we first describe two techniques for partitional clustering of hyperspectral data. Then, the segmentation scheme based on the partitional clustering is presented.

A. Clustering by ISODATA and EM

As was mentioned earlier, clustering implies a grouping of pixels in the spectral space. Let us consider the input hyperspectral image as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbf{R}^B, j = 1, 2, \dots, n\}$, where B is the number of spectral bands. Each pixel in the image is characterized by its spatial location (coordinates) and vector of spectral values. The information about spatial positions of pixels is not used in the clustering algorithms. However, it is taken into consideration during the second stage of the segmentation procedure (as will be explained in the next section).

The three principal stages of the clustering technique are as follows.

- 1) *Feature selection/extraction*: *Feature selection* consists in identifying a subset of the original features. *Feature extraction* consists in applying one or more transformations of the input features to produce new salient features. Either or both of these techniques can be applied to obtain the most effective set of features to be used in clustering.

As a pixel vector from hyperspectral image contains hundreds of spectral values, feature extraction/selection is often a required step to reduce the dimensionality of the clustering/classification problem. The most common transformations applied to these images are principal component analysis (PCA) [39], minimum noise fraction [40], and independent component analysis (ICA) [41]–[43]. Furthermore, when describing clustering algorithms, we consider, for the sake of simplicity, that \mathbf{X} is already a set of feature vectors (we also call it a *set of patterns*).

- 2) *Similarity measure*: Clustering aims at grouping pixels, so that pixels belonging to the same cluster are spectrally similar. To quantify this relationship, a similarity measure must be chosen. Proximity between pixels is usually measured by a distance function defined on pairs of spectral values. A simple distance measure like the Euclidean distance is often used to measure similarity between vectors. For some cases, other measures can be more relevant. Description of various distance measures can be found in [26], [44], and [45].
- 3) *Grouping*: In this step, pixels are grouped into clusters. Partitional clustering algorithms identify the partition that optimizes a clustering criterion (deducted from the similarity measure step).

Both ISODATA and EM are iterative optimization techniques. Thus, on each iteration i , a partition $\mathbf{Q}_1^i, \mathbf{Q}_2^i, \dots, \mathbf{Q}_C^i$ of the set \mathbf{X} into C clusters is computed, so that $\mathbf{Q}_c^i = \{\mathbf{x}_{j,c}^i \in \mathbf{R}^B, j = 1, 2, \dots, m_c^i\}$ contains the pixels belonging to the component c on the iteration i , where m_c^i is the number of pixels in \mathbf{Q}_c^i .

1) *ISODATA Algorithm*: As described in [26], the simplest and most frequently used criterion in partitional clustering is the squared-error criterion, which is the most suitable in the case of isolated and compact clusters. The squared error for a clustering Υ of a set \mathbf{X} into C clusters is defined as

$$e^2(\mathbf{X}, \Upsilon) = \sum_{c=1}^C \sum_{j=1}^{m_c} \|\mathbf{x}_{j,c} - \boldsymbol{\mu}_c\|^2 \quad (1)$$

where $\boldsymbol{\mu}_c$ is the centroid of the cluster c .

ISODATA clustering is a well-known algorithm introduced by Ball and Hall [28] which uses the squared-error criterion. It starts with a random initial partition of the pixel vectors into candidate clusters and then reassigns these vectors to clusters in such a way that the squared error (1) is reduced at each iteration, until a convergence criterion is achieved [39]. The algorithm permits splitting, merging, and deleting of clusters at each iteration in order to produce more accurate results and to mitigate dependence of results on the initialization.

The ISODATA algorithm is implemented in the ENVI software [46], where its application for hyperspectral images is straightforward. A vector of spectral values can be used as a feature vector for every pixel. When we have a reference map for the images, we can define a *minimum* number of clusters C_{\min} equal to the number of classes in the reference map and choose a *maximum* number of clusters C_{\max} superior to this value.

Several methods have used the ISODATA algorithm in remote sensing analysis. Kamagata *et al.* [47] applied the algorithm to classify multispectral IKONOS data. It was also used by Liew *et al.* [48] to classify hyperspectral Hyperion images.

2) *EM Algorithm*: While ISODATA is a deterministic clustering approach, the EM algorithm belongs to the group of statistical algorithms that assume a statistical model that characterizes the data.

The underlying assumption for the mixture resolving approach to cluster analysis (that includes the EM algorithm) is that the patterns are drawn from one or several distributions. The objective is to identify the parameters of each distribution. Most often, the individual components of the mixture density are assumed to be Gaussian. In this case, the parameters of a Gaussian mixture model have to be estimated.

The EM algorithm was proposed by Dempster *et al.* [29] to obtain iteratively a maximum likelihood estimate of the parameters of component densities from the patterns.

To cluster a hyperspectral image by the EM technique, we assume that pixels belonging to the same cluster are drawn from a multivariate Gaussian probability distribution. Each image pixel can be statistically modeled by the following probability density function:

$$p(\mathbf{x}) = \sum_{c=1}^C \omega_c \phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \quad (2)$$

where $\omega_c \in [0, 1]$ is the mixing proportion (weight) of cluster c with $\sum_{c=1}^C \omega_c = 1$ and $\phi(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate Gaussian density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$

$$\phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) = \frac{1}{(2\pi)^{B/2}} \frac{1}{|\boldsymbol{\Sigma}_c|^{1/2}} \times \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) \right\}. \quad (3)$$

The parameters of the distributions $\boldsymbol{\psi} = \{C, \omega_c, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c; c = 1, 2, \dots, C\}$ are estimated by an iterative method similar to the classification EM algorithm [49], as outlined in Algorithm 1 [31]. During the procedure of parameter estimation, pixels are assigned to the C clusters. Therefore, when the algorithm converges, the partitioning of the set of pixel vectors into C clusters is obtained.

Algorithm 1 EM clustering

Require:

- a set of n feature vectors (patterns) \mathbf{X}
- an upper bound C_{\max} on the number of clusters

Initialization (Iteration 0):

Let $C = C_{\max}$. Determine the first partition \mathbf{Q}_c^0 , $c = 1, 2, \dots, C$ of \mathbf{X} :

1. Choose randomly C patterns from the set \mathbf{X} to serve as cluster centers.

2. Assign the remaining patterns to the clusters on the basis of the nearest Euclidean distance to the cluster center.

For every iteration $i > 0$ (I iterations in total):

Parameter estimation step:

Estimate $\boldsymbol{\mu}_c^i$, $\boldsymbol{\Sigma}_c^i$, and ω_c^i for $c = 1, 2, \dots, C$ using the component-wise Maximum Likelihood estimates

$$\boldsymbol{\mu}_c^i = \frac{1}{m_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} \mathbf{x}_{j,c}^{i-1} \quad (4)$$

$$\boldsymbol{\Sigma}_c^i = \frac{1}{m_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} (\mathbf{x}_{j,c}^{i-1} - \boldsymbol{\mu}_c^i) (\mathbf{x}_{j,c}^{i-1} - \boldsymbol{\mu}_c^i)^T \quad (5)$$

$$\omega_c^i = \frac{m_c^{i-1}}{n}. \quad (6)$$

Cluster assignment step:

1. Assign each pattern in \mathbf{X} to one of the clusters according to the maximum *a posteriori* probability criteria

$$\mathbf{x}_j \in \mathbf{Q}_c^i : \Pr(c|\mathbf{x}_j) = \max_l \Pr(l|\mathbf{x}_j) \quad (7)$$

where

$$\Pr(c|\mathbf{x}_j) = \frac{\omega_c^i \phi_c(\mathbf{x}_j; \boldsymbol{\mu}_c^i, \boldsymbol{\Sigma}_c^i)}{\sum_{c=1}^C \omega_c^i \phi_c(\mathbf{x}_j; \boldsymbol{\mu}_c^i, \boldsymbol{\Sigma}_c^i)}. \quad (8)$$

2. Eliminate cluster c if m_c^i is less than the dimensionality of patterns, $c = 1, 2, \dots, C$. The patterns that belonged to the deleted clusters will be reassigned to the other clusters in the next iteration.

3. If the convergence criterion is not achieved, return to the parameter estimation step.

The total number of parameters to be estimated is $P = (B(B+1)/2 + B+1)C + 1$, where B is a dimensionality of feature vectors. If the value of B is large, P may be quite a large number. This may cause the problem of the covariance matrix singularity or inaccurate parameter estimation results. To avoid these problems, we reduce the spectral dimension of pixel vectors in hyperspectral data by averaging every AW neighboring bands, so that

$$x_{j,b}^{\text{av}} = \frac{\sum_{i=1}^{AW} x_{j,[(b-1)AW+i]}}{AW} \quad (9)$$

where $x_{j,i}$ is a value of pixel x_j in the input band i and $x_{j,b}^{\text{av}}$ is a value of pixel x_j in the output band b ; $j = 1, \dots, n$, $b = 1, \dots, B_{\text{av}}$, where $B_{\text{av}} = n/AW$.

This is a simple way of feature extraction. In previous studies, the feature extraction methods appropriate for hyperspectral image analysis, such as the PCA, the ICA, the ISOMAP and clustering-based band selection were considered [39], [42], [50], [51]. The research question to find the most effective features for the proposed method is a subject for future investigations.

Numerous authors have applied clustering using multivariate Gaussian distributions for segmentation and classification of multispectral [52]–[55] and hyperspectral images [56], [57]. In particular, Acito *et al.* [57] segmented each of the first six PCA components of the 92-band MIVIS image using 1-D Gaussian mixture models and then fused partial segmentation results.

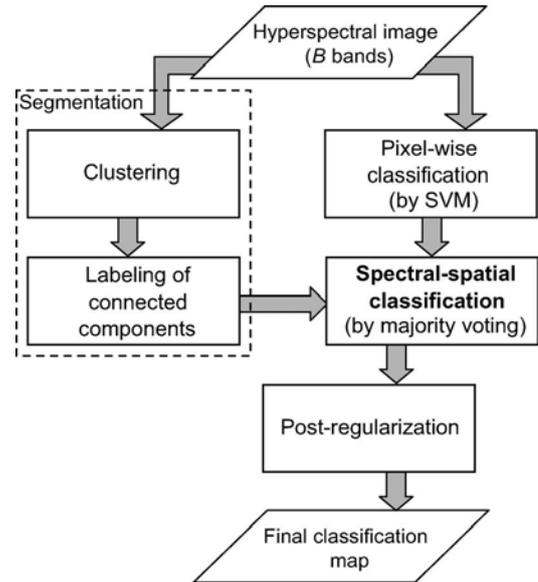


Fig. 1. Flowchart of the proposed spectral-spatial classification scheme.

Good segmentation results are reported in [57], although only visual results are presented, with no quantitative assessment. An oversegmentation effect was noted in these results, where different clusters corresponded to the same class in the ground scene.

B. Segmentation Using Clustering

The partitional clustering algorithm produces an exhaustive partitioning of the set of image pixels \mathbf{X} into C clusters. Thus, each pixel has a numerical label of the cluster it belongs to. However, as no spatial information is used during the clustering procedure, pixels with the same cluster label can be connected in the image plane, thus forming a spatial region, or they can belong to disjoint regions within the spatial coordinates. Therefore, in order to obtain a segmentation map (where each connected spatial region has a unique label), a connected-component-labeling algorithm must be applied to the output image partitioning obtained by the clustering algorithm [26], [55]. This algorithm allocates different labels for disjoint regions in the image plane that were placed in the same cluster.

If the spatial dimensions of an image are not large, a classical connected-component algorithm using the union-find data structure can be used [58]. In the case of large-sized images, such algorithms as in [59] and [60] can be applied, as well as other sequential and parallel algorithms (references can be found for example in [60]). The *Segmentation* block in Fig. 1 thus consists of two stages: *Clustering* and *Labeling of connected components*.

The obtained segmentation map can be oversegmented, as reported, for example, in [57]. However, for the research presented in this paper, oversegmentation is not a crucial problem, since the final goal is not to obtain the segmentation result but to classify the image. Thus, we are searching for spatial regions of pixels that belong to the same physical object in order to incorporate this information into a spectral-spatial classifier. From this discussion, it is evident that undersegmentation is not desired. As oversegmentation is preferable to undersegmentation,

a four-neighborhood connectivity is preferable to use while performing the labeling of connected components.

III. SPECTRAL–SPATIAL CLASSIFICATION SCHEME

The flowchart of the proposed spectral–spatial classification scheme for hyperspectral data is shown in Fig. 1.

At the input, we have a B -band hyperspectral image $\mathbf{X} = \{\mathbf{x}_j \in \mathbf{R}^B, j = 1, 2, \dots, n\}$ and a training set map.

The proposed spectral–spatial classifier is based on the *majority vote* rule. In previous studies, this approach was applied in a similar way in [61] for multispectral (four-band IKONOS) images and in [37] for hyperspectral data, giving a good performance. The approach is principally the combination of unsupervised segmentation and pixel wise classification results. The proposed method consists of the following steps (see Figs. 1 and 2).

- 1) *Segmentation*: A hyperspectral image is segmented into homogeneous regions using partitional clustering, as described in the previous section. The number of clusters (C_{\min}/C_{\max} for the ISODATA and C_{\max} for the EM) can be chosen based on the information about the considered image (i.e., how many groups of materials with similar spectra are present). C_{\min} must be chosen not to be less than the number of classes. The upper bound of classes C_{\max} can be chosen slightly superior to the number of classes. If less than C_{\max} clusters are present in the image, both algorithms have the possibility to merge clusters.
- 2) *Pixel wise classification*: Independently of the segmentation procedure, a pixel wise classification of the image is performed. We propose to use an SVM classifier with the Gaussian radial basis function (RBF) kernel for this purpose, which has given good accuracies in classification of hyperspectral data [14], [15], [17]. Parameters of the classifier can be tuned by m -fold cross validation.
- 3) *Spectral–spatial classification*: Then, for every region in the segmentation map, all the pixels are assigned to the most frequent class within this region (we call this the *majority vote* rule).

Please note that unlike in the fixed-window-based approach, the majority voting is not performed using a fixed neighborhood but using an adaptive neighborhood. For each pixel, the region it belongs to, as defined by the segmentation step, is used as its neighborhood for the majority voting on the spectral classification algorithm.

- 4) *PR*: Finally, spatial PR of the classification map is performed. The aim of this postprocessing step is to reduce the noise in the classification map after the majority vote procedure. For this purpose, the classification map is filtered, using the masks shown in Fig. 3 (that are 8- and 16-neighborhoods of a pixel, called Chamfer neighborhoods [62]). The PR is performed as follows.

a) For every pixel in the classification map: If more than T_1 neighbors in the eight-neighborhood [see Fig. 3(a)] have the class label L that is different from that of the considered pixel, assign this label L to the considered pixel. Perform this filtering until stability is reached (none of the pixels changes its label).

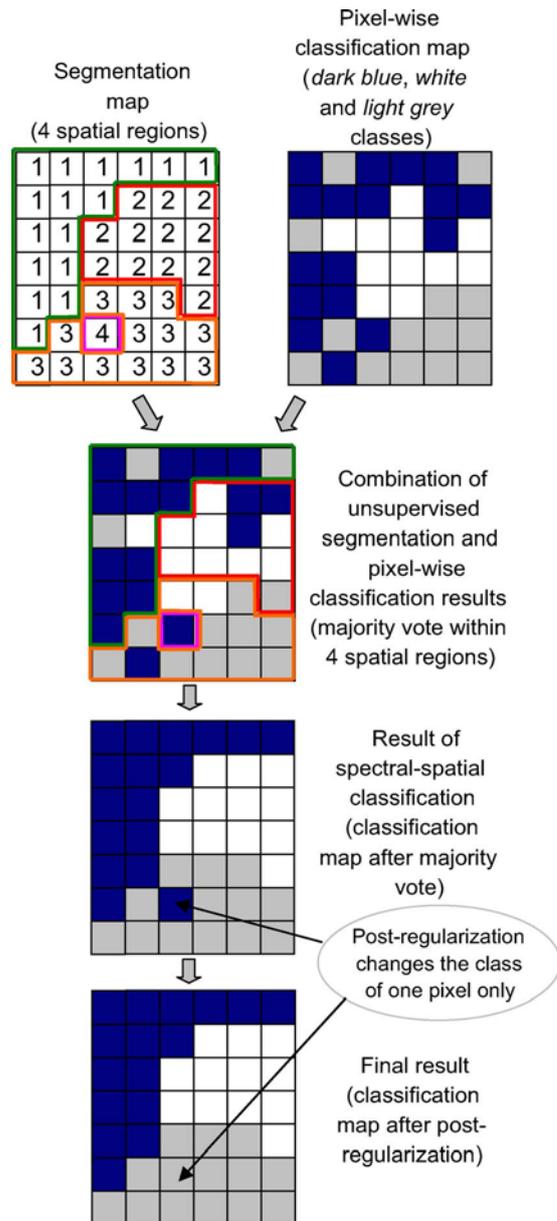


Fig. 2. Example of spectral–spatial classification.

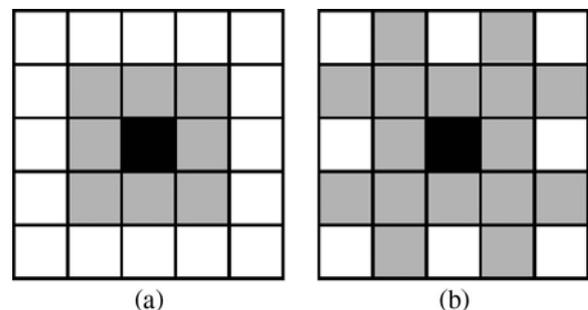


Fig. 3. Chamfer neighborhoods (in gray) for a black pixel: (a) 8 neighbors and (b) 16 neighbors.

b) For every pixel: If more than T_2 neighbors in the 16-neighborhood [see Fig. 3(b)] have the label L different from that of the considered pixel, assign the label L to the considered pixel. Perform this step until stability is reached.

- c) Repeat regularization on the eight-neighborhood (with threshold $T3$).

The threshold values $T1$ – $T3$ must be chosen to be equal or superior to a half of the number of pixels in the considered neighborhood in order to ensure the unique solution of the algorithm. The PR step results in more homogeneous regions in the classification map. However, the filtering of the classification map does not use any spectral pixel wise information. The effectiveness of this procedure depends on the sizes of the structures in the image. If the image resolution is not very high, the object in the image scene can be of the size of one or a few pixels. In this case, this object is in danger to be removed from the classification map by the PR. The filtering conditions can be restricted or relaxed by varying the threshold values $T1$ – $T3$. If $T_j(j = 1, \dots, 3)$ decreases, the regularization has a stronger effect. Hence, the results become more homogeneous. However, the risk to remove small but significant features increases.

Fig. 2 shows an example of the combination of spatial and spectral information using the proposed spectral–spatial classification method.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Hyperspectral Image Data Set

Two different data sets were used for the experiments, with different contexts (one urban area and one agricultural area), different spatial resolutions (1.3 and 20 m per pixel, respectively), and different number of bands (103 and 220 bands, respectively). These two data sets and the corresponding results are presented in the next two sections.

B. Spectral–Spatial Classification of the University of Pavia Image

The *University of Pavia* image is of an urban area that was acquired by the ROSIS-03 optical sensor over the University of Pavia, Italy (provided by Deutsches Zentrum für Luft- und Raumfahrt). The image is 610×340 pixels, with a spatial resolution of 1.3 m per pixel. The number of data channels in the original recorded image is 115 (with a spectral range from 0.43 to 0.86 μm). The 12 most noisy channels have been removed, and the remaining 103 bands were used for the experiments. The reference data contain nine classes of interest. Table II details these classes, with the number of test and training samples for each class. Fig. 4(a) shows a three-band false color image. The reference data are shown in Fig. 4(b).

First, partitional clustering of the *University of Pavia* image was performed using the two techniques described in Section II.

For the ISODATA algorithm, considering that image pixels belong to one of the nine classes, we chose the number of clusters as $C_{\min} = 9$ and $C_{\max} = 10$. The algorithm resulted in splitting all the pixels into nine clusters. A higher upper bound of the number of clusters was also tested, but in that case, the algorithm merged clusters. Furthermore, the number of regions in the resulting segmentation map increased, and the segmentation results were not improved as compared to the original initialization.

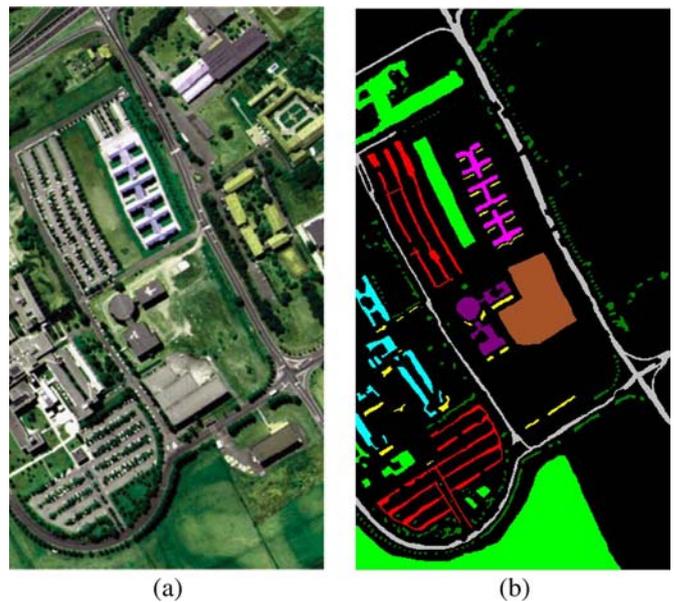


Fig. 4. *University of Pavia* image. (a) Three-band color composite. (b) Reference data: Asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks, and shadow.

The EM clustering algorithm was performed with the *maximum* number of clusters $C_{\max} = 10$. As explained in Section II, the spectral dimension needs to be reduced before applying the EM algorithm. For feature reduction, a ten-band image was obtained by averaging over every ten neighboring bands.¹ At the output of the EM algorithm, the grouping of the image pixels into ten clusters was obtained. As for the ISODATA algorithm, the increase of the upper bound of the number of classes leads to the increase of the number of regions in the segmentation map, i.e., a more severe oversegmentation.

Fig. 5(a) and (b) shows the unsupervised classification maps obtained by the ISODATA and the EM algorithms, respectively. In each of these figures, different colors correspond to different clusters (which are not associated with any physical structures, as the maps are obtained by unsupervised techniques). As shown from the figures, the main spatial structures in the scene are well defined. Based on a visual inspection, the two obtained segmentation results are of comparable accuracies. The obtained unsupervised maps are clearly oversegmented, i.e., there are cases where the regions of pixels belonging to the same object were classified to different clusters (for instance, pixels from the region of *bare soil* in the center of the image were classified into several clusters).

The classical connected-component algorithm using the union-find data structure [58] was applied to these two unsupervised classification maps (using a four-neighborhood connectivity). The resulting segmentation maps contained 20 952 and 21 450 regions for the ISODATA and the EM techniques, respectively. In both cases, some regions contain a whole single physical object. For instance, a big structure belonging to the *metal sheets* class in the center of the image is mostly represented by one region. At the same time, a lot of small,

¹The 103-band image was split into ten groups of ten bands; the three remaining bands were omitted.

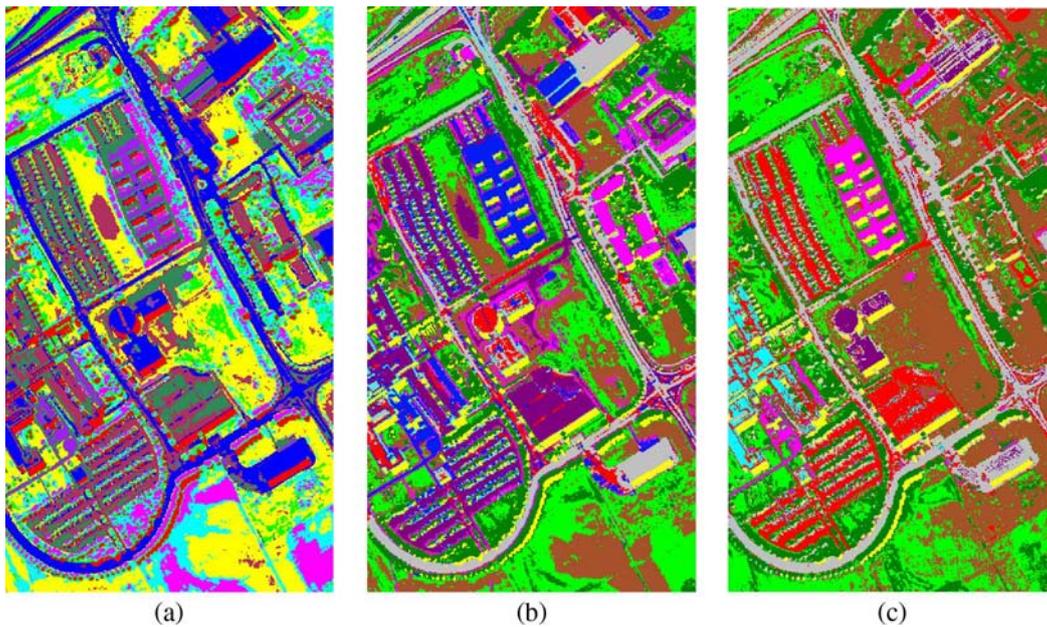


Fig. 5. Unsupervised classification maps for the *University of Pavia* image obtained by (a) ISODATA and (b) EM. (c) Supervised SVM classification map.

TABLE I
ASSESSMENT OF THE UNDERSEGMENTATION FOR THE *University of Pavia*
IMAGE (FOR REGIONS CONTAINING PIXELS OF TWO DIFFERENT
CLASSES; NUMBER OF PIXELS BELONGING TO EACH CLASS)

ISODATA		EM	
meadows	trees	meadows	trees
2	1	35	29
623	1	39	10
125	79	973	2
20	1	11	125
		1	90
asphalt	bricks		
1021	2		
trees	shadows		
1	57		

up to one pixel, regions are present, which explains the large number of regions in the obtained segmentation maps.

To assess quantitatively the accuracy of segmentation results, the analysis of undersegmentation/oversegmentation level was conducted. Two resulting segmentation maps and the image of the reference data [see Fig. 4(b)] were used for this purpose. In order to conclude if any undersegmentation is present in the considered segmentation results, images of a segmentation map and a reference data were superposed so that the reference data were partitioned into regions defined by the segmentation map. Then, the number of different classes within each region was computed. Nonlabeled pixels in the reference data were not taken into account (thus, if none of the pixels from a particular region was labeled, this region did not participate into the procedure of undersegmentation assessment).

For the undersegmentation assessment of the ISODATA segmentation map, 1560 regions were considered. Among them, 1554 regions only contained labeled pixels of the same class. Only six regions contained labeled pixels belonging to two different classes. For the EM segmentation map, among the 2029 considered regions, only 5 regions contained pixels from 2 different classes; other regions were not undersegmented. Table I gives the detailed information about undersegmented regions.

As shown from the table, the undersegmentation occurs mostly between classes *meadows* and *trees*. From this analysis, it can be concluded that the undersegmentation is almost not present in the obtained results. Therefore, the segmentation maps can, as a matter of fact, be used in the proposed spectral–spatial classification scheme.

Furthermore, in order to investigate the level of oversegmentation, we computed how many regions from the segmentation map each connected component in the reference data contained (if only a part of the region is present within the considered connected component, it was also counted as one region with this component). The reference data contained 265 connected components. Fig. 6 shows results for both ISODATA and EM segmentation maps (in a logarithmic scale). For the ISODATA (EM) segmentation results, 120 (163), 70 (56), and 25 (12) connected components from the reference data contained 1, 2, and 3 regions from the segmentation map, respectively. Thus, for both segmentation results, more than 81% of the connected components contained no more than 3 regions from the segmentation map. For several connected components, the number of regions that they contain is somewhat larger. In most cases, these components contain a large number of pixels (in Fig. 6, the total number of pixels for each connected component in reference data is also visualized). The average ratio of the number of pixels in the connected component and the number of regions within the component is equal to 16.78 and 21.46 for the ISODATA and the EM segmentation results, respectively. Based on this, it can be concluded that the oversegmentation is present in the obtained segmentation results. However, as explained in Section II, oversegmented maps of spatial regions can be used in the proposed spectral–spatial classification scheme without the risk to worsen classification accuracies obtained by the pixel wise classification.

After the initial segmentation step, the pixel wise classification step was performed using the multiclass pairwise (one versus one) SVM classifier, with the Gaussian RBF kernel

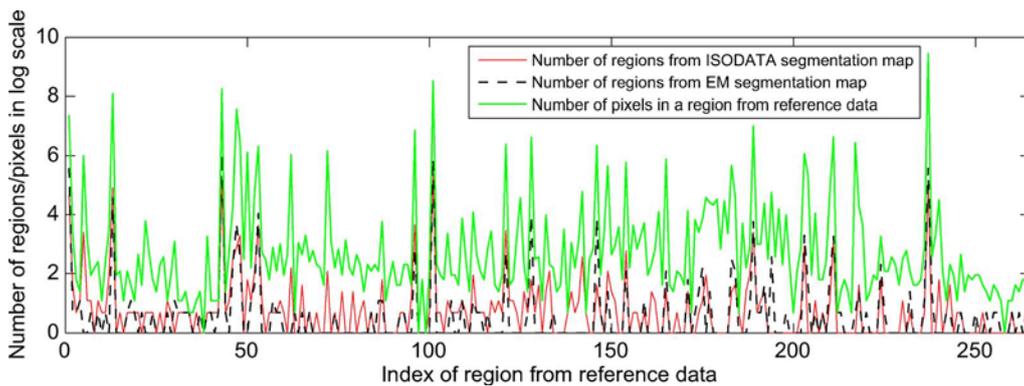


Fig. 6. Assessment of oversegmentation: Number of regions from the segmentation map contained in the connected components from the reference data and total number of pixels in each connected component from the reference data. Results are shown in a logarithmic scale.

TABLE II
INFORMATION CLASSES, TRAINING-TEST SAMPLES, AND CLASS-SPECIFIC ACCURACIES IN PERCENTAGE
FOR THE *University of Pavia* IMAGE (“PR” MEANS THE INCLUSION OF A PR STEP)

Class		Samples		SVM		SVM+ISODATA		SVM+EM		EMP
No	Name	Train	Test		PR		PR		PR	
1	asphalt	548	6304	84.93	90.13	92.16	94.40	90.72	93.45	95.36
2	meadows	540	18146	70.79	73.73	87.07	87.45	92.73	93.78	80.33
3	gravel	392	1815	67.16	68.93	61.43	61.32	82.09	82.53	87.61
4	trees	524	2912	97.77	98.76	98.59	98.63	99.21	99.38	98.37
5	metal sheets	265	1113	99.46	99.91	99.46	99.91	100	100	99.48
6	bare soil	532	4572	92.83	96.87	97.38	97.88	96.78	97.38	63.72
7	bitumen	375	981	90.42	93.68	99.69	100	92.46	94.19	98.87
8	bricks	514	3364	92.78	97.41	98.69	99.02	97.80	98.31	95.41
9	shadows	231	795	98.11	98.62	97.86	97.86	97.74	97.86	97.68

(by means of the LIBSVM library [63]). The optimal parameters C and γ were chosen by fivefold cross validation: $C = 128$ and $\gamma = 0.125$. The resulting classification map is shown in Fig. 5(c). The results of pixel wise classification were combined with the segmentation results using a *majority vote* approach (as explained in Section III).

The concluding PR step was performed on the pixel wise classification map and on two maps obtained by the spectral–spatial classification. Based on experimental results, we have chosen the threshold values $T1 = T3 = 5$ and $T2 = 12$. These values are considered as being a good tradeoff for filtering the noise while minimizing the risk of losing small but significant objects in the classification map.

Table III gives the global classification accuracies for the pixel wise SVM and the combined spectral–spatial classification before and after PR. The following measures of accuracy were used: overall accuracy (OA is the percentage of correctly classified pixels), average accuracy (AA is the mean of class-specific accuracies, i.e., the percentage of correctly classified pixels for each class), and kappa coefficient (κ , formula can be found in [39]). The class-specific accuracies are presented in Table II. Fig. 7 shows the classification maps for the pixel wise SVM and the spectral–spatial classification after the PR step. In order to compare the obtained results with previous works that used an SVM and spatial information for hyperspectral image classification, we have included in Tables II and III accuracies of mathematical morphology-based classification of the *University of Pavia* image using SVM, principal components, and extended morphological profiles (EMPs); results are taken from the work of Plaza *et al.* [64], where the same training and testing samples were used for classification. This method was

recently proposed by Benediktsson *et al.* [65] and has given good classification accuracies. Other results of spectral–spatial classification of the considered image can be found in [15], [19], and [66].

As can be seen from Table III, the SVM classifier gives high classification accuracies. The incorporation of the segmentation map obtained by clustering techniques into spectral–spatial classifier significantly improves the classification accuracies. The best global accuracies are achieved when using the spectral–spatial classifier based on the clustering by the EM algorithm with the PR step. In this case, the OA is improved by 13.7% and the AA improved by 7.0% compared to the pixel-wise SVM classification. The accuracies were substantially improved after spatial PR, with the improvement being more significant when this step is performed after a pixel wise SVM classification. This result meets expectations as the spectral–spatial classification already removes noise in the classification map, leading to more homogeneous regions. Therefore, less noise is left to be removed by means of the PR step.

The spectral–spatial classification improves the classification accuracies for almost all the classes (see Table II), except for the class *shadows*. For this class, the PR of the pixel wise SVM classification map improves the classification accuracy slightly. However, when performing the spectral–spatial classification, the classification accuracy is nonsignificantly reduced (two more pixels are misclassified compared to the results of the pixel wise classification). For the other classes, classification accuracies are improved in a range of 0.5%–23.0%. The spectral–spatial classification based on the ISODATA clustering gives the best classification accuracies for the classes *bare soil*, *bitumen*, and *bricks*, while for the classes *meadows*, *trees*, and

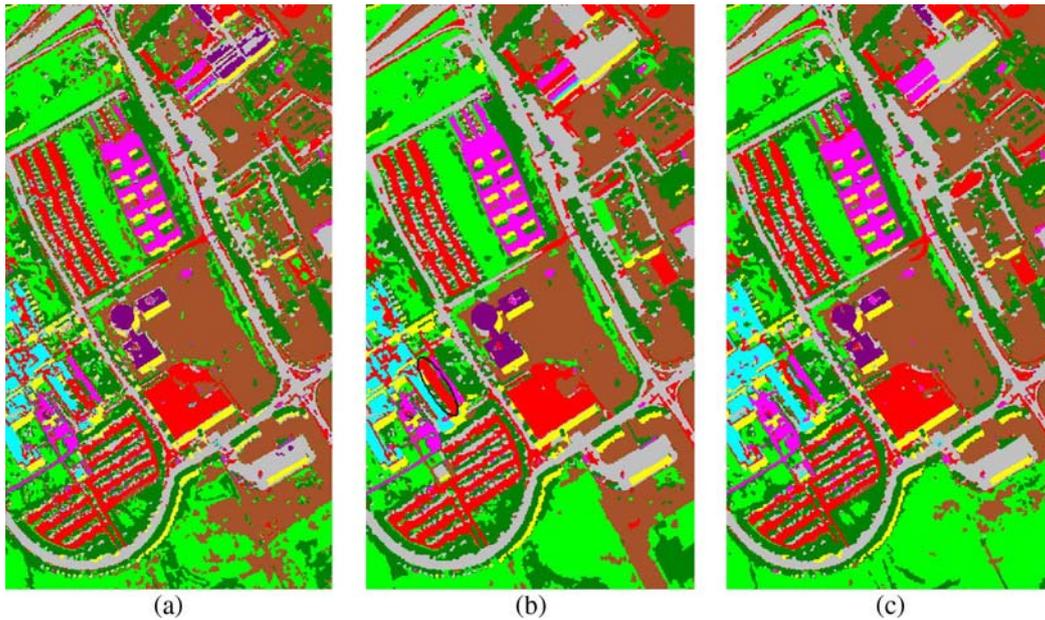


Fig. 7. Classification maps for the *University of Pavia* image after PR: (a) Pixel wise SVM classification, (b) SVM + ISODATA, and (c) SVM + EM.

TABLE III
GLOBAL CLASSIFICATION ACCURACIES IN PERCENTAGE
FOR THE *University of Pavia* IMAGE ("after PR"
MEANS THE INCLUSION OF A PR STEP)

Method		OA	AA	κ
Pixel-wise SVM		81.01	88.25	75.86
	<i>after PR</i>	84.27	90.89	79.90
SVM+ISODATA		90.57	92.48	87.62
	<i>after PR</i>	91.20	92.94	88.43
SVM+EM		93.59	94.39	91.48
	<i>after PR</i>	94.68	95.21	92.92
EMP		85.22	90.76	80.86

metal sheets, the EM clustering algorithms led to the best results. In particular, the class *meadows* is much more accurately classified when the spatial information is used (improvement of the classification accuracy by 23.0% when using the EM clustering). This class describes mostly large regions in the image. Furthermore, the incorporation of the information from the segmentation map results in much more homogeneous regions. The classes *bitumen* and *metal sheets* were identified with a 100% accuracy by the spectral–spatial classification using the ISODATA and the EM, respectively.

The use of spatial information in a classifier, by incorporating a segmentation map and performing PR, significantly reduces noise in the classification map (see Figs. 5(c) and 7). The classification maps obtained by the spectral–spatial classification before the PR are not shown as these maps are very similar to those after the PR. That is explained by the fact that majority voting within regions of the segmentation map has already removed most of the noise in the classification map, as previously mentioned. In Fig. 7(b), it can be noted that one object belonging to the class *gravel* (marked by the black ellipse in the figure) was wrongly assigned to the class *bricks*. In that case, the segmentation by ISODATA helped in identifying this object as one homogeneous object (the ISODATA identified two big regions within this object). However, as the pixel wise SVM classifier has assigned most of the pixels to the class

bricks, the whole object was assigned to this class by the *majority vote* rule. Another approach for combination of spatial and spectral information in classification could be more suitable in this case.

These results have shown that the proposed spectral–spatial classification scheme, using majority voting within the regions in the segmentation map obtained by partitioning clustering techniques, leads to improved classification accuracies and more homogeneous objects in the resulting classification maps when compared to the pixel wise classification. The approach is particularly suitable for classification of large spatial structures in the image. However, when including the spatial information from the segmentation map or from the closest neighborhoods, we risk to assimilate small structures in the image with the larger structures in their neighborhood (particularly if their spectral responses are not very different). Therefore, small structures are in danger of disappearing in the final classification map when performing the spectral–spatial classification. Accurate segmentation results help to overcome this problem.

The classification accuracies shown in this paper are higher than all previous results that we have found in the literature for this particular data set [15], [19], [64], [66]. In particular, when we compare the obtained results with the recent results of spectral–spatial classification using the SVM and EMPs (see Tables II and III), the proposed approach leads to significantly higher global accuracies and to higher class-specific accuracies for most of the classes. Thus, the segmentation using clustering, enabling the inclusion of the spatial information in a classifier, appears to be an appropriate technique for finding homogeneous objects in a hyperspectral image of an urban area.

C. Spectral–Spatial Classification of the Indiana Image

The proposed spectral–spatial classification scheme was tested on the *Indiana* image of an agricultural area, with more bands (number of bands $B = 220$) and a lower spatial resolution, as compared to the *University of Pavia* image.

TABLE IV
INFORMATION CLASSES, NUMBER OF LABELED SAMPLES, AND CLASS-SPECIFIC ACCURACIES
IN PERCENTAGE FOR THE *Indiana* IMAGE (“PR” MEANS THE INCLUSION OF A PR STEP)

Class		No. of samples in reference data	SVM		SVM+ISODATA		SVM+EM	
No	Name			PR		PR		PR
1	Corn-no till	1434	74.59	83.27	79.32	80.48	71.65	71.42
2	Corn-min till	834	64.58	78.30	84.95	88.02	84.15	86.68
3	Corn	234	58.77	81.99	75.83	85.31	60.66	65.40
4	Soybeans-no till	968	69.76	78.81	83.85	84.19	82.02	82.47
5	Soybeans-min till	2468	79.21	93.88	93.16	95.77	95.05	97.12
6	Soybeans-clean till	614	75.41	89.69	85.17	89.87	90.05	92.59
7	Alfalfa	54	32.65	36.73	12.24	6.12	0	0
8	Grass/pasture	497	87.05	90.18	93.08	93.75	93.97	95.31
9	Grass/trees	747	92.72	97.62	94.80	96.88	99.11	99.26
10	Grass/pasture-mowed	26	29.17	29.17	91.67	91.67	0	0
11	Hay-windrowed	489	96.37	97.96	97.51	97.51	99.09	99.09
12	Oats	20	22.22	22.22	16.67	11.11	0	0
13	Wheat	212	90.58	97.38	93.19	98.43	98.95	98.95
14	Woods	1294	91.07	95.88	97.17	97.85	95.36	96.82
15	Bldg-Grass-Tree-Drives	380	65.50	78.36	79.53	85.38	69.30	79.24
16	Stone-steel towers	95	84.88	84.88	86.05	87.21	86.05	86.05

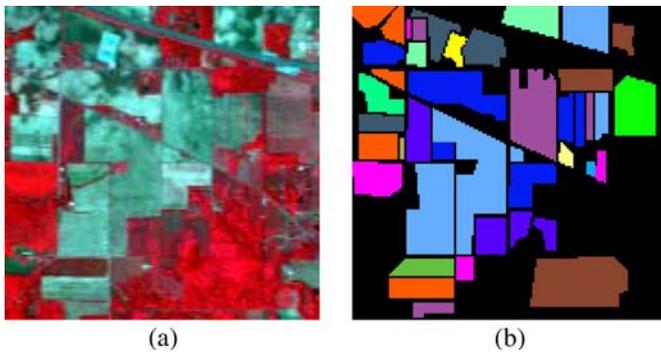


Fig. 8. *Indiana* image. (a) Three-band color composite (bands 50, 27, and 17). (b) Reference data: Corn-no till, corn-min till, corn, soybeans-no till, soybeans-min till, soybeans-clean till, alfalfa, grass/pasture, grass/trees, grass/pasture-mowed, hay-windrowed, oats, wheat, woods, bldg-grass-tree-drives, and stone-steel towers.

The *Indiana* image was recorded by the AVIRIS sensor over the Indian Pines test site in Northwestern Indiana [38]. The image has a spatial dimension of 145×145 pixels, and the spatial resolution is 20 m per pixel. The full spectral range of 220 channels was used for the experiments. Sixteen classes of interest are considered, which represent mostly different types of crops and are detailed in Table IV, with a number of samples for each class in the reference data. A three-band false color image and the reference data are shown in Fig. 8. We have chosen randomly 10% of the samples for each class from the reference data as training samples, and the remaining samples composed the test set.

Both the ISODATA and the EM algorithm were applied to perform a partitional clustering of the image. The ISODATA was performed with $C_{\min} = 16$ and $C_{\max} = 19$. The EM clustering was applied to the 22-band image (obtained by averaging over every 10 neighboring bands of the original data) with $C_{\max} = 17$. Both algorithms grouped the image pixels into 17 clusters.

The unsupervised classification maps obtained by the ISODATA and the EM algorithm are shown in Fig. 9(a) and (b), respectively (different colors correspond to different clusters). As in the previous experiment, the images are oversegmented,

but here, the spatial structures, corresponding to the crop fields, can be recognized. We can also see [when comparing these classification maps with the reference data in Fig. 8(b)] that some pixels belonging to different classes are classified by the clustering techniques to the same cluster. For instance, in Fig. 9(a), at the center of the image, there are two large light-green regions of pixels that belong to the same cluster. These regions represent different crop fields: *soybeans-no till* (class 4, violet color in the reference data) and *soybeans-min till* (class 5, light-blue color in the reference data). The spectral responses of the pixels from these two classes are similar, and the clustering algorithms group them to the same cluster. However, as we are interested in obtaining a segmentation map, where the image is partitioned into regions, without any additional information about the region, it is important that these two regions of pixels belonging to the same cluster are disconnected in space.

To obtain segmentation maps, connected-component labeling of the unsupervised classification maps was performed using the same algorithm as for the previous data set and four-neighborhood connectivity. The resulting segmentation maps for the ISODATA and the EM techniques contained 3977 and 3728 regions, respectively. As explained for the *University of Pavia* image, the segmentation using clustering produces a map with comparatively large regions along with a lot of very small and one-pixel regions.

Multiclass, one versus one, SVM classification was performed on the original image using the Gaussian RBF kernel. The parameters C and γ were determined by fivefold cross validation, which gave $C = 1024$ and $\gamma = 2^{-7}$. Fig. 9(c) shows the obtained classification map. After the pixel wise SVM classification, majority voting within the regions from each of the segmentation maps was performed. Then, PR was applied to the two classification maps obtained by the spectral-spatial classification and to the pixel wise classification map (with $T1 = T3 = 5$ and $T2 = 12$).

Tables IV and V give the class-specific and the global classification accuracies, respectively, for the pixel wise and the spectral-spatial classification, without and with the PR step. The classification maps for the pixel wise and the spectral-spatial classification after the PR are shown in Fig. 10.

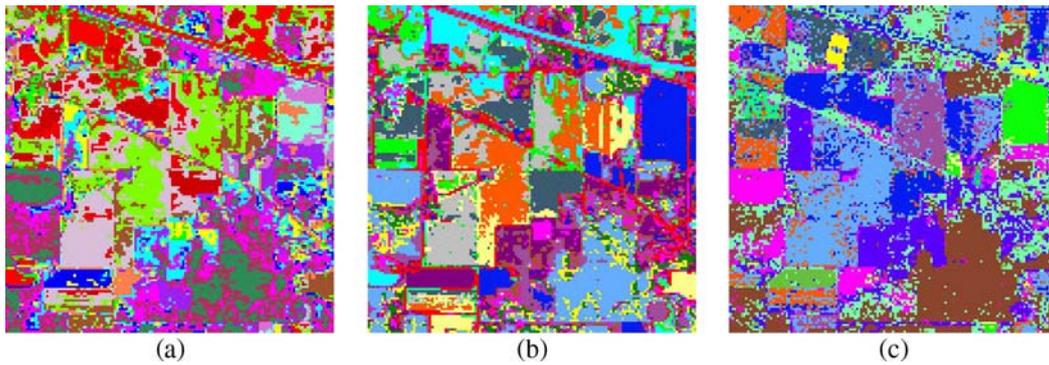


Fig. 9. Unsupervised classification maps for the *Indiana* image obtained by (a) ISODATA and (b) EM. (c) Supervised SVM classification map.

TABLE V

GLOBAL CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE *Indiana* IMAGE (“after PR” MEANS THE INCLUSION OF A PR STEP)

Method		OA	AA	κ
Pixel-wise SVM		78.76	69.66	75.73
	after PR	88.58	77.27	86.93
SVM+ISODATA		88.53	79.01	86.90
	after PR	90.64	80.60	89.31
SVM+EM		87.25	70.34	85.43
	after PR	88.83	71.90	87.24

Here, it is worth mentioning that the low spatial resolution of the *Indiana* image leads to the presence of highly mixed pixels. This makes the tasks of the unsupervised segmentation by clustering and the supervised SVM classification more complicated. We previously discussed an example where pixels belonging to different classes were grouped to the same cluster. One of the reasons of this inaccuracy in the unsupervised classification could be the presence of spectrally mixed pixels in the image.

Another complication of the segmentation and classification tasks is caused by a significant difference in the number of pixels in the image and in the reference data for different classes, which varies in the reference data from 20 to 2468 pixels per class. Some classes represent big crop fields, while others represent very small fields.

The 10% of samples for each class were chosen randomly from the reference data as a training set for the SVM classifier. Therefore, some classes were represented by a few samples in the training set (two samples only for the class *oats*), which probably do not provide a fair-enough representation of the class.

Despite these complications, the SVM classifier correctly classified 78.76% of pixels from the test set. The classification accuracies for the classes *alfalfa*, *grass/pasture-mowed*, and *oats*, which were represented by only a few samples in the training set (further called as *small classes*), are low (less than 33%). Consequently, the average classification accuracy is only 69.66%.

As can be seen from Tables IV and V, the simple filtering (PR) improves the classification accuracies significantly. The OA and AA are improved by 9.8% and 7.6%, respectively, after the PR step of the pixel wise SVM classification. Accuracies of almost all classes are significantly improved, except for some *small classes* (for them, the accuracies are not changed, but for the class *alfalfa*, the accuracy is improved). This is explained by the fact that most of the classes in the image represent large

crop fields, and the simple filtering makes these regions of fields homogeneous, thereby improving the classification accuracies.

The best global accuracies are obtained when performing the spectral–spatial classification using ISODATA clustering and the PR. In that case, the OA and AA are improved by 11.9% and 10.9%, respectively, compared to the pixel wise SVM classification. Almost all the class-specific accuracies are improved. When looking at the results for the *small classes*, the classification accuracy was significantly improved for the class *grass/pasture-mowed* (from 29.17% to 91.67%), while for the classes *alfalfa* and *oats*, accuracies are reduced (the problem of the spectral–spatial classification for the *small classes* will be discussed hereinafter).

The OA for the spectral–spatial classification using the EM clustering is slightly lower than that when using the ISODATA clustering technique. The EM clustering approach led to the best classification results for some classes (six classes, as can be seen from Table IV). However, for the *small classes* (*alfalfa*, *grass/pasture-mowed*, and *oats*), none of the pixels from the test set was identified correctly by this classifier, and that reduced the average classification accuracy. The potential misclassification of *small classes* is actually caused by the low spatial resolution of image, the presence of classes with similar spectral responses, and the small number of samples per class in the image/training set. The two main reasons for the problem of the classification of these classes can be defined.

- 1) Very small crop fields of *grass/pasture-mowed* and *oats* were assimilated with their neighboring regions (which represented the big fields of *grass/pasture* and *grass/trees*, respectively) when performing the segmentation and the majority voting.
- 2) For the class *alfalfa*, the EM clustering grouped the pixels from the *alfalfa* and the *hay-windrowed* (a big light-green field in the right part of the image) fields into the same cluster. And the pixel wise SVM classifier assigned the majority of the *alfalfa* pixels from the test set to the class *hay-windrowed*, as the spectral responses of these two classes were similar [see Figs. 9(b) and (c) and 10(c)]. The segmentation map contains a separate region that corresponds to the *alfalfa* field, but according to the *majority vote* rule, all the pixels were assigned to *hay-windrowed*, which is an incorrect class.

As mentioned before, one of the problems with the partitional clustering techniques concerns the dependence of the results on

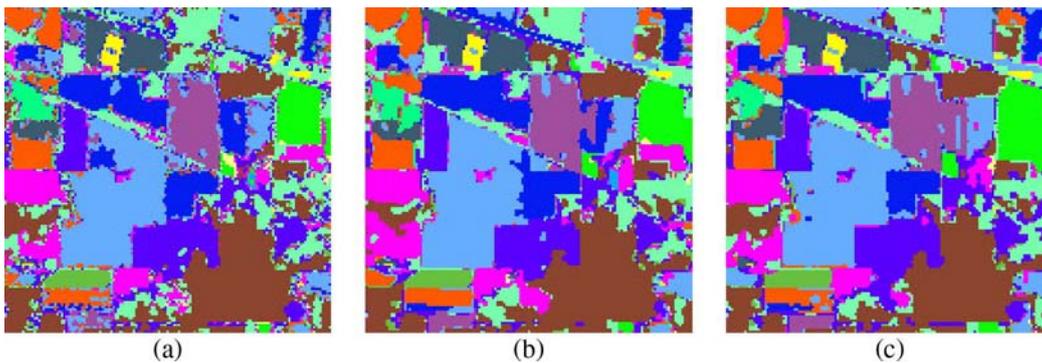


Fig. 10. Classification maps for the *Indiana* image after PR. (a) Pixel wise SVM classification, (b) SVM + ISODATA, and (c) SVM + EM.

TABLE VI
PROCESSING TIME IN SECONDS FOR CLUSTERING ALGORITHMS AND
SVM CLASSIFICATION FOR THE *University of Pavia* IMAGE
(B —DIMENSIONALITY OF PATTERNS;
 I —NUMBER OF ITERATIONS)

Algorithm	B	I	time
ISODATA	103	8	185
EM	10	3	3
SVM (training+classification)	103	-	3339

the initialization. For the two considered algorithms, C cluster centers are initially chosen randomly from a set of pixels, and the remaining pixels are assigned to the cluster with the closest center. If one class contains very few pixels compared to the other classes, the probability is low that one of its pixels will be chosen at the start as a cluster center. Then, when centers are recomputed and pixels are reassigned, pixels from this *small class* can compose a separate class only if their spectral response is very different from that of the other classes. However, otherwise, if there is a class with a large number of pixels that have a similar spectral response, pixels from the *small class* will probably be grouped to the same cluster with the *large class*. Furthermore, the considered clustering algorithms have mechanisms to delete clusters, based on the number of pixels in the cluster. For instance, the EM algorithm eliminates a cluster if the number of pixels in this cluster is less than the dimensionality of the pixels (as the covariance matrix of every cluster must be computed at every iteration). It can also be an obstacle to identify the pixels from a *small class* to a separate cluster (if this cluster is very small, it can be eliminated). All these reasons led to either 0% or low classification accuracies for the *small classes* when classifying the *Indiana* image.

The experimental results on the *Indiana* image (see Figs. 9(c) and 10) have confirmed that the proposed spectral–spatial classification method based on the partitional clustering results in a classification map with more homogeneous regions when compared to pixel wise classification. The proposed scheme is particularly suitable for classification of images with large spatial structures. Furthermore, it is also suitable if different classes have dissimilar spectral responses and a comparable number of pixels (of the same order).

More classification results for the *Indiana* image can be found in [64] and [67] for comparison. The accuracies in the referenced works are not directly compared with those given in this paper because different training–testing sets are used. How-

ever, it can be concluded that our approach performs well compared to other previously proposed classification approaches.

D. Consideration of Computational Complexity for the Spectral–Spatial Classification Method

When comparing the results of two classifiers, an important issue is the computational complexity and the processing time of each classifier. Although the pixel wise SVM classifier gives good classification accuracies, it is a computationally demanding algorithm for high-dimensional data and/or when the number of training samples is large [19], [68]. The training part of the SVM classification is the most time-consuming, in particular the tuning of parameters by cross validation. We conducted experiences on an Intel Core 2 Duo 2.40-GHz processor with 3.5-GB RAM. The processing times for the training and classification parts of the *University of Pavia* image by means of the LIBSVM library were 3240 and 99 s, respectively. However, in recent works, methods and parallel implementations to speed up the SVM training and classification have been proposed [64], [69], [70], [71].

When we perform a segmentation of an image by clustering and combine spatial information with the results of pixel wise classification, the processing time obviously increases, when compared to pixel wise classification only. However, the partitional clustering algorithms are much less time-consuming than the SVM classification algorithm. The computational complexity of both the ISODATA and EM clustering algorithms is $O(nCB^2I)$, where I is the number of iterations (until the convergence of algorithm). Therefore, we can say that the processing time depends mainly on the dimensions of the image. Table VI summarizes the processing time for clustering algorithms versus SVM classification for the *University of Pavia* image as a function of dimensionality of patterns and number of iterations. These results are not directly comparable, as different software packages were used (ENVI software to apply ISODATA and the C++ implementation for the EM algorithm). However, it can be seen that the EM algorithm ran much faster than the ISODATA algorithm mainly because of a lower spectral dimensionality. Furthermore, the processing time for the used clustering techniques is significantly smaller than the time for the SVM classification (although, in general, this ratio depends on the number of training samples and the clustering algorithm). In addition, as was mentioned before, efficient implementations of the clustering algorithms are possible.

In addition, it must be mentioned that the segmentation and pixel wise classification of an image can be executed at the same time on different processing units (since none of these tasks depends on the results of another task). In that case, after the pixel wise classification is completed, majority voting and PR are applied, which are both very fast algorithms. Furthermore, as shown earlier, the incorporation of spatial information significantly improves accuracies. When the spatial information is incorporated, the classification in the pixel wise classification step can be speeded up by sacrificing some percentage of the classification accuracy. That can be achieved either by decreasing the number of pixels in the training set or by using another less time-consuming classifier than the SVM for pixel wise classification.

V. CONCLUSION

A new spectral-spatial classification scheme for hyperspectral images was presented. The proposed method combines the results of a pixel wise SVM classification and a segmentation map obtained by partitional clustering. This is achieved by performing a majority voting on the pixel wise spectral classification using adaptive neighborhoods defined by the segmentation map. The use of both the ISODATA and the Gaussian mixture resolving techniques for hyperspectral image segmentation was investigated. The incorporation of spatial information from the segmentation in the classifier produces a classification map with more homogeneous regions, as compared to only pixel wise classification of hyperspectral data. Here, the remaining noise in the classification map was further reduced by a fixed-window-based postfiltering.

Experimental results have shown that the proposed method improves the classification accuracies and provides classification maps with more homogeneous regions when compared to pixel wise classification.

The developed scheme is particularly suitable for classification of images with large spatial structures, when spectral responses of the different classes are dissimilar and the classes contain a comparable number of pixels. The drawback of the proposed method is that when including spatial information from the segmentation map or from the closest neighborhoods in a classifier, small spatial structures face a risk of being assimilated with larger neighboring structures if the spectral responses are not significantly different.

In the future, we will investigate the use of feature extraction to find the most effective features to be used in the clustering. In particular, applying feature reduction transformations enables the reduction of the spectral dimension while the most important information for classification is preserved. That may lead to a better distinction between classes and thus to better segmentation results.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments that helped improve this paper significantly. The authors would also like to thank Deutsches Zentrum für Luft- und Raumfahrt, P. Gamba from the University of Pavia, Italy, and D. Landgrebe from Purdue University,

USA, for providing the hyperspectral data. The authors would also like to thank I. Kåsen, T. V. Haavardsholm, and T. Skauli from FFI, Norway, for a fruitful discussion about the EM algorithm.

REFERENCES

- [1] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. S. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, Sep. 1998.
- [2] L. Rickard, W. Basedow, E. Zalewski, P. Silverglate, and M. Landers, "HYDICE: An airborne system for hyperspectral imaging," in *Proc. SPIE Imaging Spectrometry Terrestrial Environ.*, 1993, vol. 1937, pp. 173–179.
- [3] B. Stevenson, R. O'Connor, W. Kendall, A. Stocker, W. Schaff, R. Holasek, D. Even, D. Alexa, J. Salvador, M. Eismann, R. Mack, P. Kee, S. Harris, B. Karch, and J. Kershenstein, "The civil air patrol ARCHER hyperspectral sensor system," in *Proc. SPIE Airborne ISR Syst. Appl. II*, 2005, vol. 5787, pp. 17–28.
- [4] T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields, "The HyMap airborne hyperspectral sensor: The system, calibration and performance," in *Proc. 1st EARSEL Workshop Imaging Spectroscopy*, Oct. 1998, pp. 37–43.
- [5] J. Pearlman, C. Segal, L. Liao, S. Carman, M. Folkman, B. Browne, L. Ong, and S. Ungar, "Development and operations of the EO-1 Hyperion imaging spectrometer," in *Proc. SPIE Earth Observing Syst. V*, 2000, vol. 4135, pp. 243–253.
- [6] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Dordrecht, The Netherlands: Kluwer, 2003.
- [7] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. New York: Wiley, 2003.
- [8] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agric.*, vol. 39, no. 2, pp. 67–93, 2003.
- [9] H. Zhou, Z. Mao, and D. Wang, "Classification of coastal areas by airborne hyperspectral image," in *Proc. SPIE Opt. Technol. Atmos., Ocean, Environ. Stud.*, May 2005, vol. 5832, pp. 471–476.
- [10] S. Subramanian, N. Gat, M. Sheffield, J. Barhen, and N. Toomarian, "Methodology for hyperspectral image classification using novel neural network," in *Proc. SPIE Algorithms Multispectral Hyperspectral Imagery III*, A. E. Iverson, S. S. Shen, Eds., Aug. 1997, vol. 3071, pp. 128–137.
- [11] H. Yang, F. V. D. Meer, W. Bakker, and Z. J. Tan, "A back-propagation neural network for mineralogical mapping from AVIRIS data," *Int. J. Remote Sens.*, vol. 20, no. 1, pp. 97–110, 1999.
- [12] C. Hernández-Espinosa, M. Fernández-Redondo, and J. Torres-Sospedra, "Some experiments with ensembles of neural networks for classification of hyperspectral images," in *Proc. ISNN*, vol. 1, 2004, pp. 912–917.
- [13] C. Vaiphasa, "Innovative genetic algorithm for hyperspectral image classification," in *Proc. Int. Conf. Map Asia*, 2003, p. 20.
- [14] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [15] M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data," Ph.D. dissertation, Grenoble Inst. Technol., Grenoble, France, 2007.
- [16] J. A. Gualtieri and R. F. Crompt, "Support vector machines for hyperspectral remote sensing classification," *Proc. SPIE*, vol. 3584, pp. 221–232, Jan. 1998.
- [17] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Evaluation of kernels for multiclass classification of hyperspectral remote sensing data," in *Proc. ICASSP*, May 2006, pp. II-813–II-816.
- [18] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.
- [19] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 3804–3814, Oct. 2008.
- [20] A. Farag, R. Mohamed, and A. El-Baz, "A unified framework for map estimation in remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, Jul. 2005.
- [21] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*, ser. McGraw-Hill Series in Computer Science. New York: McGraw-Hill, 1995.

- [22] K. Fu and J. Mui, "A survey on image segmentation," *Pattern Recognit.*, vol. 13, no. 1, pp. 3–16, 1981.
- [23] P. Lambert and L. Macaire, "Filtering and segmentation: The specificity of color images," in *Proc. Int. Conf. Color Graph. Image Process.*, 2000, vol. 1, pp. 57–64.
- [24] R. Haralick and L. Shapiro, "Survey: Image segmentation techniques," *Comput. Vis. Graph. Image Process.*, vol. 29, no. 1, pp. 100–132, 1985.
- [25] P. Paclík, R. Duin, G. V. Kempen, and R. Kohlus, "Segmentation of multi-spectral images using the combined classifier approach," *Image Vis. Comput.*, vol. 21, no. 6, pp. 473–482, Jun. 2003.
- [26] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
- [27] S. Lee and M. Crawford, "Unsupervised multistage image classification using hierarchical clustering with a Bayesian similarity measure," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 312–320, Mar. 2005.
- [28] G. Ball and D. Hall, "ISODATA, a novel method of data analysis and classification," Stanford Univ., Stanford, CA, Tech. Rep. AD-699616, 1965.
- [29] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [30] N. B. Venkateswarlu and P. S. V. S. K. Raju, "Fast ISODATA clustering algorithms," *Pattern Recognit.*, vol. 25, no. 3, pp. 335–342, Mar. 1992.
- [31] Y. Tarabalka, T. V. Haavardsholm, I. Kåsen, and T. Skauli, "Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing," *J. Real-Time Image Process.*, pp. 1–14, 2008. DOI:10.1007/s11554-008-0105-x.
- [32] F. Lafarge, X. Descombes, and J. Zerubia, "Textural kernel for SVM classification in remote sensing: Application to forest fire detection and urban area extraction," in *Proc. ICIP*, Sep. 2005, vol. 3, pp. III-1096–III-1099.
- [33] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [34] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.
- [35] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Adaptive pixel neighborhood definition for the classification of hyperspectral images with Support Vector Machines and composite kernel," in *Proc. ICIP*, San Diego, CA, 2008, pp. 1884–1887.
- [36] S. van der Linden, A. Janz, B. Waske, M. Eiden, and P. Hostert, "Classifying segmented hyperspectral data from a heterogeneous urban environment using support vector machines," *J. Appl. Remote Sens.*, vol. 1, no. 1, p. 013 543, Oct. 2007.
- [37] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in *Proc. IGARSS*, Boston, MA, 2008, pp. III-652–III-655.
- [38] AVIRIS NW Indiana's Indian Pines 1992 Data Set. [Online]. Available: ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/92AV3C (original files) and ftp://ftp.ecn.purdue.edu/biehl/PC_MultiSpec/ThyFiles.zip (ground truth)
- [39] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. New York: Springer-Verlag, 1999.
- [40] 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," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
- [41] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.
- [42] J. Wang and C.-I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1586–1600, Jun. 2006.
- [43] S. Moussaoui, H. Hauksdottir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Douté, and J. A. Benediktsson, "On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation," *Neurocomputing*, vol. 71, no. 10–12, pp. 2194–2208, Jun. 2008.
- [44] E. Diday and J. C. Simon, "Clustering analysis," in *Digital Pattern Recognition*. Berlin, Germany: Springer-Verlag, 1976, pp. 47–94.
- [45] G. Noyel, J. Angulo, and D. Jeulin, "Morphological segmentation of hyperspectral images," *Image Anal. Stereol.*, vol. 26, pp. 101–109, 2007.
- [46] ITT Corporation. For further information about the ENVI software. [Online]. Available: http://rsinc.com/envi/
- [47] N. Kamagata, K. Hara, M. Mori, Y. Akamatsu, Y. Li, and Y. Hoshino, "A new method of vegetation mapping by object-based classification using high resolution satellite data," in *Proc. 1st Int. Conf. Object-Based Image Anal.*, 2006, vol. XXXVI-4/C42.
- [48] S. C. Liew, C. W. Chang, and K. H. Lim, "Hyperspectral land cover classification of EO-1 Hyperion data by principal component analysis and pixel unmixing," in *Proc. IGARSS*, 2002, vol. 6, pp. 3111–3113.
- [49] G. Celeux and G. Govaert, "A classification EM algorithm for clustering and two stochastic versions," *Comput. Stat. Data Anal.*, vol. 14, no. 3, pp. 315–332, Oct. 1992.
- [50] X. R. Wang, S. Kumar, F. Ramos, and T. Kaupp, "Probabilistic classification of hyperspectral images by learning nonlinear dimensionality reduction mapping," in *Proc. Int. Conf. Inf. Fusion*, Jul. 2006, pp. 1–8.
- [51] A. Martinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 4158–4171, Dec. 2007.
- [52] P. Masson and W. Pieczynski, "SEM algorithm and unsupervised segmentation of satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp. 618–633, May 1993.
- [53] A. Solberg, T. Taxt, and A. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.
- [54] S. Pal and P. Mitra, "Multispectral image segmentation using the roughset-initialized EM algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2495–2501, Nov. 2002.
- [55] J. Driesen and P. Scheunders, "A multicomponent image segmentation framework," in *Proc. ACTVS*, 2008, pp. 589–600.
- [56] S. Beaven, D. Stein, and L. Hoff, "Comparison of Gaussian mixture and linear mixture models for classification of hyperspectral data," in *Proc. IGARSS*, 2000, vol. 4, pp. 1597–1599.
- [57] N. Acito, G. Corsini, and M. Diani, "An unsupervised algorithm for hyperspectral image segmentation based on the Gaussian mixture model," in *Proc. IGARSS*, Jul. 2003, vol. 6, pp. 3745–3747.
- [58] L. Shapiro and G. Stockman, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [59] J.-O. Lapeyre and R. Strandh, "An efficient union-find algorithm for extracting the connected components of a large-sized image," Lab. Bordelais de Recherche en Inf., Bordeaux, France, Jan. 2004. Tech. Rep.
- [60] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Comput. Vis. Image Underst.*, vol. 89, no. 1, pp. 1–23, Jan. 2003.
- [61] A. Widayati, B. Verbist, and A. Meijerink, "Application of combined pixel-based and spatial-based approaches for improved mixed vegetation classification using IKONOS," in *Proc. 23th Asian Conf. Remote Sens.*, 2002, pp. 1–8.
- [62] G. Borgefors, "Distance transformations in digital images," *Comput. Vis. Graph. Image Process.*, vol. 34, no. 3, pp. 344–371, Jun. 1986.
- [63] C. Chang and C. Linin *LIBSVM—A Library for Support Vector Machines*, 2008. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm
- [64] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, in press. DOI: 10.1016/j.rse.2007.07.028.
- [65] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
- [66] S. Aksoy, "Spatial techniques for image classification," in *Signal and Image Processing for Remote Sensing*, C. H. Chen, Ed. New York: Taylor & Francis, 2006, pp. 491–513.
- [67] S. Tadjudin and D. Landgrebe, "Classification of high dimensional data with limited training samples," Ph.D. dissertation, School Elect. Comput. Eng., Purdue Univ., Hammond, IN, May 1998.
- [68] N. Cristianini and J. Shawe-Taylor, *Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [69] H. P. Graf, E. Cosatto, L. Bottou, I. Dourdanovic, and V. Vapnik, "Parallel support vector machines: The cascade SVM," in *Advances in Neural Information Processing Systems*, vol. 17. Cambridge, MA: MIT Press, 2005.

- [70] T. Habib, J. Inglada, G. Mercier, and J. Chanussot, "Speeding up support vector machine (SVM) image classification by a kernel series expansion," in *Proc. IEEE ICIP*, Oct. 2008, pp. 865–868.
- [71] B. C. Catanzaro, N. Sundaram, and K. Keutzer, "Fast Support Vector Machine training and classification on graphics processors," EECSS Dept., Univ. California, Berkeley, Berkeley, CA, Tech. Rep. UCB/EECS-2008-11, 2008.



Yuliya Tarabalka (S'08) received the B.S. degree in computer science from the Ternopil Ivan Pul'uj State Technical University, Ternopil, Ukraine, in 2005, and the M.Sc. degree in signal and image processing from the Grenoble Institute of Technology (INPG), Grenoble, France, in 2007. Since 2007, she has been working toward the Ph.D. degree, a cojoint degree between the University of Iceland, Reykjavik, Iceland, and INPG.

From July 2007 to January 2008, she was a Researcher with the Norwegian Defence Research Establishment, Norway. She is currently conducting her research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab, INPG). Her current research work is funded by the "HYPER-I-NET" Marie Curie Research Training Network. Her research interests are in the areas of image processing, pattern recognition, hyperspectral imaging, and development of efficient algorithms.



Jón Atli Benediktsson (S'84–M'90–SM'99–F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively.

He is currently a Pro Rector for academic affairs and a Professor of electrical and computer engineering with the University of Iceland. He has held visiting positions with the Department of Information and Communication Technology, University of Trento,

Trento, Italy, since 2002; the School of Computing and Information Systems, Kingston University, Kingston upon Thames, U.K., during 1999–2004; the Joint Research Centre of the European Commission, Ispra, Italy, in 1998; the Technical University of Denmark, Lyngby, Denmark, in 1998; and the School of Electrical and Computer Engineering, Purdue University, in 1995. In August 1997, he was a Fellow with the Australian Defence Force Academy, Canberra, A.C.T., Australia. From 1999 to 2004, he was the Chairman of the energy company Metan Ltd. He was also the Chairman of the University of Iceland's Science and Research Committee from 1999 to 2005. Since 2006, he has been the Chairman of the University of Iceland's Quality Assurance Committee. He coedited (with Prof. D. A. Landgrebe) a Special Issue on Data Fusion of TGARS (May 1999) and (with P. Gamba and G. G. Wilkinson) a Special Issue on Urban Remote Sensing from Satellite (October 2003). His research interests include remote sensing, pattern recognition, neural networks, image processing, and signal processing, and he has published extensively in those fields.

Dr. Benediktsson was the Past Editor and is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGARS) and an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. From 1999 to 2002, prior to being Editor, he was an Associate Editor for TGARS. From 1996 to 1999, he was the Chairman of GRSS Technical Committee on Data Fusion and was elected to the Administrative Committee of the GRSS for the terms 2000–2002 and 2008–2010. In 2002, he was appointed as Vice President of Technical Activities of GRSS, and in 2008, he was appointed as Vice President of Professional Activities of GRSS. He was the founding Chairman of the IEEE Iceland Section and served as its Chairman from 2000 to 2003. He was a member of a NATO Advisory Panel of the Physical and Engineering Science and Technology Subprogram (2002–2003). He was a member of Iceland's Science and Technology Council (2003–2006) and a member of the Nordic Research Policy Council (2004). He received the Stevan J. Kristof Award from Purdue University, in 1991, as an outstanding graduate student in remote sensing. In 1997, he was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award. In 2000, he was granted the IEEE Third Millennium Medal. In 2004, he was a corecipient of the University of Iceland's Technology Innovation Award. In 2006, he received the yearly research award from the Engineering Research Institute, University of Iceland. In 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a member of Societas Scinetiarum Islandica and Tau Beta Pi.



Jocelyn Chanussot (M'04–SM'04) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA-French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab, INPG). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007) and a member of the IEEE Geoscience and Remote Sensing AdCom (2009–2011). He is the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Cochair of the GRS Data Fusion Technical Committee (2005–2008) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and for *Pattern Recognition* (2006–2008). He is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Paper 4

Y. Tarabalka, J. Chanussot, J. A. Benediktsson, “Classification based marker selection for watershed transform of hyperspectral images,” in Proc. of *IGARSS'09*, Cape Town, South Africa, 2009, pp. III-105 - III-108.

CLASSIFICATION BASED MARKER SELECTION FOR WATERSHED TRANSFORM OF HYPERSPECTRAL IMAGES

Yuliya Tarabalka^{1,2}, Jocelyn Chanussot¹, and Jón Atli Benediktsson²

¹GIPSA-Lab, Grenoble Institute of Technology, France

²University of Iceland, Reykjavik, Iceland

e-mail: yuliya.tarabalka@hyperinet.eu, jocelyn.chanussot@gipsa-lab.inpg.fr, benedikt@hi.is

ABSTRACT

A new method for segmentation and classification of hyperspectral images is proposed. The method is based on a pixel-wise classification followed by selection of the most reliable classified pixels as markers for watershed segmentation. Furthermore, each marker defined from classification results is associated with a class label. By assigning the class label of each marker to all the pixels within the region grown from this marker, a spectral-spatial classification map is obtained. Experimental results are presented on a 200-band AVIRIS image of the Northwestern Indiana's Indian Pine site. The developed segmentation and classification scheme significantly decreases oversegmentation, improves classification accuracies and provides classification maps with more homogeneous regions, when compared to pixel-wise classification or previously proposed spectral-spatial classification techniques.

Index Terms— hyperspectral images, segmentation, classification, watershed, marker selection

1. INTRODUCTION

Hyperspectral (HS) imaging technology, which acquires hundreds of spectral channels, opens new perspectives in classification of remote sensing images. An extensive literature is available on the classification of HS images. Among them, one can separate pixel-wise processing techniques that work on the spectral information only (one of the most frequently used techniques are Support Vector Machines (SVM) [1]) and spectral-spatial classification techniques that take into consideration both the *spectra* of the pixels and their spatial context [2].

In our previous works, we have proposed to define spatial structures in a HS image by performing segmentation and considering every region from a segmentation map as an adaptive homogeneous neighborhood for all the pixels within this region [3, 4]. In particular, the extension of watershed segmentation to HS was investigated in [3]. Typically, the

standard watershed transform results in a severe oversegmentation (every local *minimum* of the gradient leads to one region). One of the ways to cope with this problem consists in performing a marker-controlled watershed segmentation [5]. This approach determines markers for each region of interest (each object in the image) and transforms the gradient image in such a way that the region markers are the only local *minima* of the resulting image.

In this paper, we propose to **determine markers for a watershed** on a HS image by using results of a pixel-wise classification. Thus, a **new segmentation and classification scheme** for HS data is proposed. The objectives of the proposed method are:

1. To decrease the oversegmentation and thus improve the segmentation results.
2. Each classification-based marker is associated with a class label. Therefore, the corresponding class can be assigned to every region in the segmentation map, resulting in a spectral-spatial classification map.

The outline of this paper is as follows. In the next section, a marker-controlled watershed segmentation and classification scheme is described. Section 3 discusses experimental results. Finally, conclusions are drawn in Section 4.

2. MARKER-CONTROLLED WATERSHED SEGMENTATION AND CLASSIFICATION SCHEME

Let us consider the input HS image as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbf{R}^B, j = 1, 2, \dots, n\}$, where B is the number of spectral channels. For the classification, each pixel must be assigned to one of the K classes of interest. The flow-chart of the proposed segmentation and classification scheme is shown in Figure 1. In the following, each step is described.

2.1. Pixel-wise classification

The first step consists in performing a pixel-wise classification of the HS image. We propose to use an SVM classifier for this purpose which is extremely well suited to classify HS

This research is supported in part by the Marie Curie Research Training Network "HYPER-I-NET".

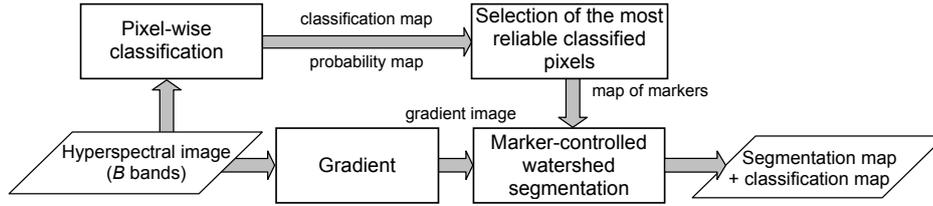


Fig. 1. Flow-chart of the proposed segmentation and classification scheme.

data [1, 2]. This step results in a classification map (where each pixel contains a label of the class it belongs to) and a probability map (if a pixel is assigned to the class k , the probability map contains a probability estimate for this pixel to actually belong to the class k).

We propose to compute probability estimates for multi-class classification by pairwise coupling, using the method implemented in LIBSVM library [6]. The objective is to estimate, for each pixel \mathbf{x} , the probabilities to belong to each class of interest:

$$p_k = p(y = k|\mathbf{x}), k = 1, \dots, K. \quad (1)$$

For this purpose, first pairwise class probabilities $r_{ij} \approx p(y = i|y = i \text{ or } j, \mathbf{x})$ are estimated [6], and then probabilities (1) are computed, by solving linear systems. Finally, a probability map is constructed, by assigning to each pixel the *maximum* probability estimate $\max(p_k), k = 1, \dots, K$.

2.2. Selection of the most reliable classified pixels

The aim of this step is to choose the most reliable classified pixels, in order to define suitable markers.

The problem of automatic marker selection was previously discussed in the literature, mostly for grey-scale and color images. Markers are often defined by searching flat zones, zones of homogeneous texture or image *extrema* [5]. Noyel *et al.* [7] performed unsupervised classification of a HS image, and selected pertinent clusters to be the markers. The author's objective was to segment a specific structure (glue occlusions). In our study, the objective is to mark (select a marker for) each spatial object in the image.

We propose to use probability estimates obtained as a result of the pixel-wise classification for this purpose, in order to keep the most reliable classified pixels as markers. A simple way of marker selection consists in thresholding the probability map. In other words, if the probability of the considered pixel to belong to the assigned class k is higher than a given threshold, this pixel is selected to join the markers. In the resulting map of markers, each marker pixel is associated with the class defined by the pixel-wise classifier. However, this technique has the following disadvantage: when using

markers for a watershed segmentation, each marker leads to one region in the segmentation map. Therefore, we need as many markers as the desired number of regions. However, if classes k_i and k_j are spectrally similar, pixels belonging to one of these classes have quasi-equal probability to belong to each of them. From here, these classified pixels are not reliable. Therefore, we risk to lose the regions corresponding to either class k_i or k_j in the final segmentation map. This leads to *undersegmentation*, which is highly undesired.

To mitigate this problem, we propose the following procedure. First, a connected components labeling is performed on the pixel-wise classification map. Then, each connected region is analysed as follows:

- If a region is large enough, it should contain a marker. It is determined as the $P\%$ of pixels within the connected component with the highest probability estimates.
- If a region is small, it should lead to a marker only if it is very reliable; potential marker is formed by the pixels with probability estimates higher than a defined threshold.

As a conclusion, each connected set of pixels with the same class in the classification map provides either one or zero marker. One should stress that a marker is not necessarily a connected set of pixels: It can be spatially split into several subsets.

2.3. Gradient

Independently of the previous steps, a gradient of the HS image is computed. A one-band gradient is needed as an input for the watershed segmentation. Approaches to compute a one-band gradient from the HS image are analyzed in [7, 3].

2.4. Marker-controlled watershed segmentation

In our previous work, the use of the watershed segmentation for HS images was investigated [3]. The algorithm of Vincent and Soille [8] has been used to perform a watershed.

In this study, we want to incorporate the use of markers into the watershed algorithm. One of the ways to do it is

described in [5]. First, the marker image is created:

$$f_m(\mathbf{x}) = \begin{cases} 0, & \text{if pixel } \mathbf{x} \text{ belongs to a marker,} \\ t_{max}, & \text{otherwise.} \end{cases} \quad (2)$$

Then, the *minima* imposition technique is applied to the gradient image f_g . The resulting image f_{gmi} is computed as a morphological reconstruction by erosion of $(f_g + 1) \wedge f_m$ (point-wise *minimum* between the gradient image and the marker image) from the marker image f_m :

$$f_{gmi} = R_{(f_g+1) \wedge f_m}^{\varepsilon}(f_m) \quad (3)$$

Once this image is computed, any watershed algorithm can be applied on the image f_{gmi} . We use the described technique of *minima* imposition. However, several issues must be taken into consideration. When creating a map of markers, it must be checked that each marker is spatially disconnected from any other marker (markers with different class labels must not “touch” each other).

The algorithm of Vincent and Soille [8] for watershed creates a new region label for each local *minimum*, and then grows regions from *minima*. In our case, a marker can contain non-adjacent groups of pixels. After the *minima* imposition, this will lead to several local *minima* in the filtered gradient image. We aim to obtain a segmentation map where each marker leads to one single region. Furthermore, we are searching for a segmentation map where each pixel belongs to one of the regions defined by the markers, without border (watershed) pixels between adjacent regions. In order to obtain such a map, we propose the following scheme:

1. After *minima* imposition, the Vincent and Soille algorithm is applied on the filtered gradient image f_{gmi} .
2. In the resulting segmentation map, each pixel contains the label of the region it belongs to, or the watershed pixel label. Here, region labels do not correspond to marker labels. Each marker leads to one or several regions. In order to obtain a segmentation map without border pixels, each watershed pixel is assigned to the neighboring region with the “closest” median, *i.e.*, the distance between the vector median of this region and the watershed pixel vector is minimal [3].
3. Regions belonging to the same marker are merged together and are associated with this marker.

Finally, pixels of each region are assigned to the class of its marker. This results in a spectral-spatial classification map.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed spectral-spatial classification scheme was tested on the *Indiana* image acquired by the AVIRIS sensor over the Indian Pines test site in Northwestern Indiana.

The image is 145 by 145 pixels, the spatial resolution is 20 m per pixel. Twenty water absorption bands have been removed, and a 200-band image was used for experiments. Sixteen classes of interest are considered. More information about the image and reference data can be found in [4]. We have chosen randomly 50 samples for each class from the reference data as training samples, except for classes “*alfalfa*”, “*grass/pasture-mowed*” and “*oats*”. These classes contain a small number of samples in the reference data. Therefore, only 15 samples for each of these classes were chosen randomly to be used as training samples.

A pixel-wise classification on the 200-band image was performed, using the multi-class one *versus* one SVM classifier with the Gaussian Radial Basis Function kernel (with parameters $C = 128$, $\gamma = 2^{-6}$). Then, a map of markers was created. For this purpose, labeling of connected components on the pixel-wise classification map was performed, using 8-neighborhood connectivity. For each connected component:

- If it contained more than $M = 20$ pixels, $P = 5\%$ of its pixels with the highest probability estimates were selected as the marker for this component.
- Otherwise, a potential marker of the region was formed by the pixels with probability estimates higher than a threshold T .

In order to define the threshold T , the probability estimates for the whole image were sorted, and T was chosen equal to the lowest probability within the highest 2% of all probability estimates. From 2250 connected components in the classification map, 107 markers were selected (see Figure 2(a)). In the obtained map of markers, no marker was spatially adjacent to any other marker. Therefore, no additional processing was required in order to use the map for a *minima* imposition procedure followed by the watershed segmentation.

In order to compute a gradient on the original image, a Robust Color Morphological Gradient (RCMG) was used [3]. Then, *minima* imposition technique was performed, followed by watershed segmentation, assignment of watershed pixels to adjacent regions and merging of regions corresponding to the same marker (as described in Section 2). The resulting segmentation map is almost not oversegmented (watershed without markers led to 1277 regions, while marker-based watershed resulted in 107 regions). Finally, pixels of each region were assigned to the class of its marker. The resulting spectral-spatial classification map is shown in Figure 2(b). Classification accuracies for the pixel-wise SVM and the proposed spectral-spatial classification (*M-WHED*) are given in Table 1. In order to compare the efficiency of the proposed method with previously proposed techniques, we have included results of classification using majority vote within adaptive neighborhoods defined by watershed segmentation map (*WHED+MV*), as described in [3].

Table 1. Classification Accuracies in Percentage for the *Indiana* Image: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ) and Class-Specific Accuracies.

	SVM	M-WHED	WHED+MV
OA	78.17	85.99	86.63
AA	85.97	86.95	91.61
κ	75.33	83.98	84.83
Corn-no till	78.18	80.35	94.22
Corn-min till	69.64	71.94	78.06
Corn	91.85	73.37	88.59
Soybeans-no till	82.03	98.91	96.30
Soybeans-min till	58.95	80.48	68.82
Soybeans-clean till	87.94	84.75	90.78
Alfalfa	74.36	94.87	94.87
Grass/pasture	92.17	95.30	95.08
Grass/trees	91.68	92.97	97.99
Grass/pasture-mowed	100	100	100
Hay-windrowed	97.72	99.54	99.54
Oats	100	100	100
Wheat	98.77	99.38	99.38
Woods	93.01	99.36	97.11
Bldg-Grass-Tree-Drives	61.52	55.45	69.39
Stone-steel towers	97.78	64.44	95.56

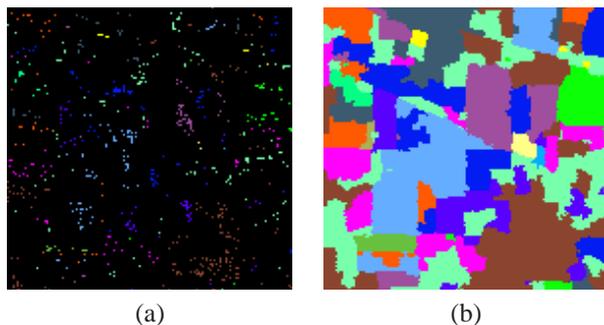


Fig. 2. (a) Map of markers and (b) classification map obtained by the scheme proposed in this paper for the *Indiana* image.

As can be seen from Table 1, both the global and most of the class-specific accuracies are improved when including spatial information into the classification. For a marker-based watershed segmentation and a majority vote scheme, overall accuracies are improved by 7.8 and by 8.6 percentage points, respectively, when compared to pixel-wise classification. The previously proposed majority vote scheme leads to the highest global classification accuracies. However, spectral-spatial classification proposed in this paper gives the highest classification accuracies for 9 of the 16 classes. The two spectral-spatial classification techniques provide classification maps with more homogeneous regions, when compared to pixel-wise classification.

4. CONCLUSIONS

In this paper, a new method for segmentation and classification of HS images is proposed. The method is based on a pixel-wise classification followed by selection of the most reliable classified pixels as markers. The markers are further used for a watershed segmentation, resulting in segmentation and spectral-spatial classification maps. The developed segmentation and classification scheme significantly decreases oversegmentation, improves classification accuracies and provides classification maps with more homogeneous regions, when compared to pixel-wise classification or previously proposed spectral-spatial classification methods.

In the future, we plan to extend the proposed scheme by analysis of regions in a segmentation map, in order to assign the region to the correct class and thus to further improve classification accuracies.

5. REFERENCES

- [1] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geos. and Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, June 2005.
- [2] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geos. and Remote Sens.*, vol. 46, no. 10, Oct. 2008.
- [3] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in *Proc. of IGARSS '08*, Boston, USA, 2008, pp. III-652 – III-655.
- [4] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *in press, IEEE Trans. Geos. and Remote Sens.*
- [5] P. Soille, *Morphological Image Analysis*, 2nd ed. Springer-Verlag, 2003.
- [6] C. Chang and C. Lin, "LIBSVM: A library for Support Vector Machines," Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.
- [7] G. Noyel, J. Angulo, and D. Jeulin, "Morphological segmentation of hyperspectral images," *Image Analysis & Stereology*, vol. 26, pp. 101–109, 2007.
- [8] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis and Machine Intel.*, vol. 13, no. 6, pp. 583–598, June 1991.

Paper 5

Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, “Segmentation and classification of hyperspectral images using Minimum Spanning Forest grown from automatically selected markers,” *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, DOI 10.1109/TSMCB.2009.2037132.

Segmentation and Classification of Hyperspectral Images Using Minimum Spanning Forest Grown From Automatically Selected Markers

Yuliya Tarabalka, *Student Member, IEEE*, Jocelyn Chanussot, *Senior Member, IEEE*, and Jón Atli Benediktsson, *Fellow, IEEE*

Abstract—A new method for segmentation and classification of hyperspectral images is proposed. The method is based on the construction of a minimum spanning forest (MSF) from region markers. Markers are defined automatically from classification results. For this purpose, pixelwise classification is performed, and the most reliable classified pixels are chosen as markers. Each classification-derived marker is associated with a class label. Each tree in the MSF grown from a marker forms a region in the segmentation map. By assigning a class of each marker to all the pixels within the region grown from this marker, a spectral-spatial classification map is obtained. Furthermore, the classification map is refined using the results of a pixelwise classification and a majority voting within the spatially connected regions. Experimental results are presented for three hyperspectral airborne images. The use of different dissimilarity measures for the construction of the MSF is investigated. The proposed scheme improves classification accuracies, when compared to previously proposed classification techniques, and provides accurate segmentation and classification maps.

Index Terms—Classification, hyperspectral images, marker selection, minimum spanning forest (MSF), segmentation.

I. INTRODUCTION

IMAGE CLASSIFICATION, which can be defined as identification of objects in a scene captured by a vision system, is one of the important tasks of a robotic system. On the one side, the procedure of accurate object identification is known to be more difficult for computers than for people [1]. On the other side, recently developed image acquisition systems (for instance, radar, lidar, and hyperspectral imaging technologies) capture more data from the image scene than a human vision system. Therefore, efficient processing systems must be developed in order to use these data for accurate image classification.

Manuscript received May 29, 2009; revised September 4, 2009. This work was supported in part by the Marie Curie Research Training Network “Hyper-I-Net.” This paper was recommended by Associate Editor D. Goldgof.

Y. Tarabalka is with the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA Lab), Grenoble Institute of Technology (INPG), 38402 Grenoble, France, and the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: yuliya.tarabalka@hyperinet.eu).

J. Chanussot is with the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA Lab), Grenoble Institute of Technology (INPG), 38402 Grenoble, France (e-mail: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr).

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSMCB.2009.2037132

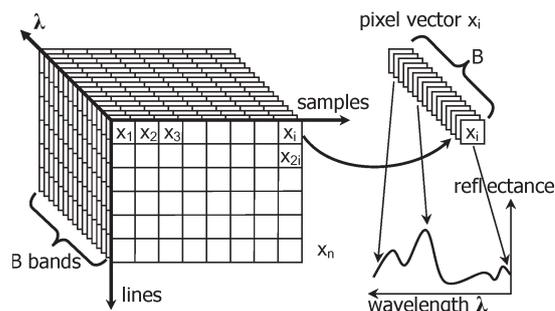


Fig. 1. Structure of a hyperspectral image.

Hyperspectral imagery records a detailed spectrum of light arriving at each pixel [2]. Hyperspectral sensors measure the energy of the received light in tens or hundreds of narrow spectral bands (data channels) in each spatial position of the image (Fig. 1 shows the structure of a hyperspectral image). This rich information per pixel increases the capability to distinguish materials and objects and thus opens new perspectives for image classification. However, a large number of spectral channels, usually coupled with limited availability of reference data,¹ present challenges to image analysis. While pixelwise classification techniques process each pixel independently without considering the information about spatial structures [3]–[5], further improvement of classification results can be achieved by considering spatial dependences between pixels, i.e., by performing spectral-spatial classification [6]–[11].

Segmentation is an exhaustive partitioning of the input image into homogeneous regions [12]. Segmentation techniques are a powerful tool to define spatial dependences. In previous works, we have performed unsupervised segmentation of hyperspectral images in order to define spatial structures [9], [13], [14]. Watershed, partitional clustering, and hierarchical segmentation techniques have been used for this purpose. Segmentation and pixelwise classification were performed independently, and then, the results were combined using a majority voting rule. Thus, every region from a segmentation map was considered as an adaptive homogeneous neighborhood for all the pixels within this region. The described technique led to a significant improvement of classification accuracies and provided more homogeneous (less noisy) classification maps when compared

¹By *reference* data, we mean manually labeled pixels which are used for training classifiers followed by assessment of classification accuracies.

to classification techniques using local neighborhoods in order to include spatial information into a classifier.

However, unsupervised image segmentation is a challenging task. Segmentation aims at dividing an image into homogeneous regions, but the measure of homogeneity is image dependent [12]. Depending on this measure, the process can result in undersegmentation (several regions are detected as one) or oversegmentation (one region is detected as several ones) of the image. In previous works [13], [14], we preferred oversegmentation to undersegmentation in order not to miss objects in the classification map. In this work, we aim to reduce oversegmentation and thus further improve segmentation and classification results. This can be achieved by using markers or region seeds [12], [15]. In previous studies, a *marker* (an internal marker) was defined as a connected component belonging to the image and associated with an object of interest [12], [15]–[17]. In our study, we define a marker as a set of image pixels (not necessarily connected; it can be composed of several spatially disjoint subsets of adjacent pixels) which is associated with one object in the image scene.

The problem of automatic marker selection has previously been discussed in the literature, mostly for gray-scale and color images. Markers are often defined by searching flat zones (i.e., connected components of pixels of constant gray-level value), zones of homogeneous texture, or image extrema [15]. Gómez *et al.* [18] applied histogram analysis to obtain a set of representative pixel values, and the markers were generated with all the image pixels with representative gray values. Jalba *et al.* [16] used connected operators filtering on the gradient image in order to select markers for a gray-scale diatom image. Noyel *et al.* [17], [19] performed classification of the hyperspectral image (using different techniques, such as Clara [20] and linear discriminant analysis) and then filtered the classification maps class by class, using morphological operators, in order to select large spatial regions as markers. Furthermore, the authors proposed to use random balls (connected sets of pixels of randomly selected sizes) extracted from these large regions as markers. In the discussed studies [16], [17], [19], the objective was to segment specific structures (blood cells, diatoms, glue occlusions, and cancerous growth).

In our study, the objective is to *mark* (select a marker for) each significant spatial object in the image. Here, by significant, we mean an object of at least one-pixel size that belongs to one of the classes of interest. As remote sensing images contain small and complex structures, automatic selection of markers is an especially challenging task.

In this paper, a new *scheme for marker-based segmentation and classification of hyperspectral images* is proposed. In particular, we propose to perform a probabilistic pixelwise classification first in order to choose the most reliable classified pixels as markers of spatial regions [21]. Furthermore, image pixels are grouped into a minimum spanning forest (MSF) [22], where each tree is rooted on a classification-derived marker. The decision to connect the pixel, which is not yet in the forest, to one of the trees in the forest is based on its similarity to one of the adjacent pixels already belonging to the forest. By assigning a class of the marker to all the pixels within the region grown from the considered marker, a spectral-

spatial classification map is obtained. Furthermore, the classification map is refined using the results of a pixelwise classification and a majority voting within the spatially connected regions [14].

The construction of an MSF belongs to graph-based approaches for image segmentation [22]–[25]. They introduce the Gestalt principles of perceptual grouping to the field of computer vision. The image is associated with a graph, the vertices of which correspond to the image entities (pixels or regions) and the edges correspond to relations between these entities. A weight associated with each edge indicates the (dis)similarity between two entities (pixels or regions). Morris *et al.* [23] have proposed to perform a graph-based image segmentation into R regions by constructing a shortest spanning tree on the image graph and then removing the $R - 1$ edges with the highest weight. Furthermore, several graph-cut-based algorithms have been developed for image segmentation [24], [25]. However, these methods perform unsupervised segmentation by splitting at each iteration one region into two subregions. This approach is fundamentally different from the work described in this paper. Several recent publications describe the use of an MSF rooted on markers for image segmentation [22], [26], [27]. However, the authors of these works do not investigate the problem of automatic marker selection. Their segmentation is based on markers provided by the user.

The proposed procedure of defining markers for each spatial object from probabilistic classification results and of building a spectral-spatial classification map for hyperspectral images by constructing an MSF rooted on classification-derived markers is a major contribution of this paper. Please note that, while, in previous studies, markers were used as seeds for image segmentation, in this paper, we introduce a new concept of the automatic marker-based spectral-spatial classification.

- 1) Markers are derived from probabilistic pixelwise classification results.
- 2) Each marker can be composed of several spatially disjoint subsets of adjacent pixels, and each marker has a class label.
- 3) By performing a region growing from the classification-derived markers, a spectral-spatial classification map is obtained.

Although the classification scheme proposed in this paper has been designed for hyperspectral data, the method is general and can successfully be applied for other types of data as well. Experimental results are demonstrated on hyperspectral airborne images recorded by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over Northwestern Indiana's Indiana Pine site and over the region surrounding the volcano Hekla in Iceland, and the image acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) over the University of Pavia in Italy.

The outline of this paper is as follows. In the next section, a classification scheme based on an MSF rooted on markers is presented. Experimental results are discussed in Section III. Finally, conclusions are drawn in Section IV.

II. SEGMENTATION AND CLASSIFICATION SCHEME

The flowchart of the proposed segmentation and classification method is shown in Fig. 2. On the input, a B-band hyperspectral image is given, which can be considered as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. Classification consists in assigning each pixel to one of the K classes of interest. In the following, each step of the proposed procedure is described.

A. Pixelwise Classification

The first step consists in performing a probabilistic pixelwise classification of the hyperspectral image. We propose to use a support vector machine (SVM) classifier [28] for this purpose. Other classifiers could be used. However, SVMs are extremely well suited to classify hyperspectral data [5], [29], [30]. We refer the reader to [5] and [28] for details on SVMs. The outputs of this step are the following:

- 1) classification map, containing class labels for each pixel;
- 2) probability map, containing probability estimates for each pixel to belong to the assigned class.

Two techniques for computing probability estimates for multiclass classification by pairwise coupling are described in [31]. We propose to use one of these methods, which is implemented in the LIBSVM library [32]. The objective is to estimate, for each pixel \mathbf{x} , the probabilities to belong to each class of interest

$$\mathbf{p} = \{p_k = p(y = k|\mathbf{x}), k = 1, \dots, K\}. \quad (1)$$

For this purpose, first, pairwise class probabilities $r_{ij} \approx p(y = i|y = i \text{ or } j, \mathbf{x})$ are estimated using an improved implementation [33] of [34]

$$r_{ij} \approx \frac{1}{1 + e^{A\hat{f} + B}} \quad (2)$$

where A and B are estimated by minimizing the negative log-likelihood function using known training data and decision values \hat{f} . Furthermore, the probabilities in (1) are computed by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{p}} \sum_{i=1}^K \sum_{j:j \neq i} (r_{ji}p_i - r_{ij}p_j)^2 \\ \text{subject to } \sum_{i=1}^K p_i = 1, \quad p_i \geq 0 \forall i. \end{aligned} \quad (3)$$

This problem has a unique solution and can be solved by a simple linear system, as described in [31]. Finally, a probability map is constructed by assigning to each pixel the *maximum* probability estimate $\max(p_k), k = 1, \dots, K$.

B. Selection of the Most Reliable Classified Pixels

The aim of this step is to choose the most reliable classified pixels in order to define suitable markers. We propose to use probability estimates obtained as a result of the pixelwise classification for this purpose in order to keep the most reliable

classified pixels as markers. A simple way of marker selection consists in thresholding the probability map. In other words, if the probability of the considered pixel belonging to the assigned class k is higher than a given threshold, this pixel is selected to join the markers. In the resulting map of markers, each marker pixel is associated with the class defined by the pixelwise classifier. The marker pixels form connected components in the map of markers so that each connected component represents one marker. The main advantage of this technique of marker selection is its simplicity. However, this method has the following disadvantage: Each marker leads to one region in the segmentation map. Therefore, we need as many markers as the desired number of regions. However, if classes k_i and k_j are spectrally similar, pixels belonging to one of those classes have a quasi-equal probability to belong to each of them. From here, these classified pixels are not reliable. Therefore, we risk to lose the regions corresponding to either class k_i or k_j in the final segmentation map. This leads to *undersegmentation*, which is highly undesired.

To mitigate this problem, we propose the following method of marker selection [see the flowchart in Fig. 3(a)].

- 1) Perform a *connected-component labeling* of the pixelwise classification map. For this purpose, a classical connected-component algorithm using the union-find data structure can be used [35].
- 2) Analyze each *connected region* as follows.
 - If a region is *large* enough, it *should contain a marker*, which is determined as $P\%$ of the pixels within the connected component with the highest probability estimates.
 - If a region is *small*, it should lead to a marker *only if it is very reliable*; a potential marker is formed by pixels with probability estimates higher than a defined threshold.

The proposed procedure is deduced from the following analysis: Based on the results of our previous studies [9], [13], [14], it is common that almost no undersegmentation is present in a pixelwise classification map. Therefore, each connected spatial region from the classification map is analyzed if it corresponds most probably to the spatial structure or if it is rather a classification noise [see the illustrative example in Fig. 3(b)]. If the size of the component is large enough to consider it as a relevant region, the most reliable pixels within this region are selected as its marker. If a component contains only a few pixels, it is investigated if these pixels were classified to a particular class with a high probability. If this is the case, the considered connected component represents a small spatial structure. Thus, a marker associated with this region should be defined. Otherwise, the component is the consequence of classification noise, and we tend to eliminate it. Therefore, no marker within this component is selected. When performing labeling of connected components for a pixelwise classification map, we propose to use an eight-neighborhood connectivity.

For the proposed marker selection procedure, the following parameters must be chosen.

- 1) A parameter M defining if a region is considered as being large or small. We propose to use a number of pixels in

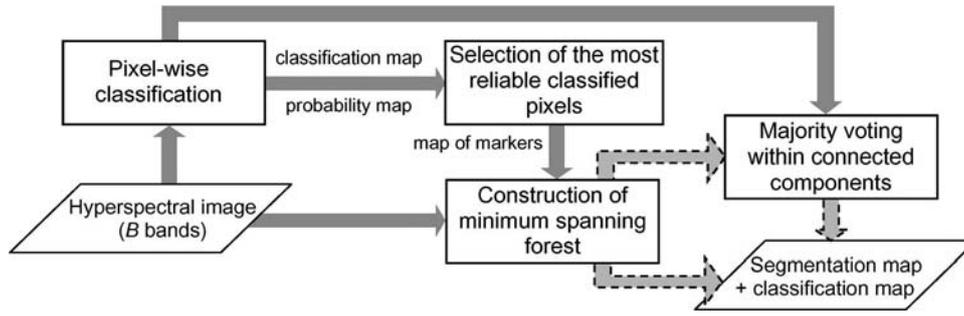


Fig. 2. Flowchart of the proposed segmentation and classification scheme.

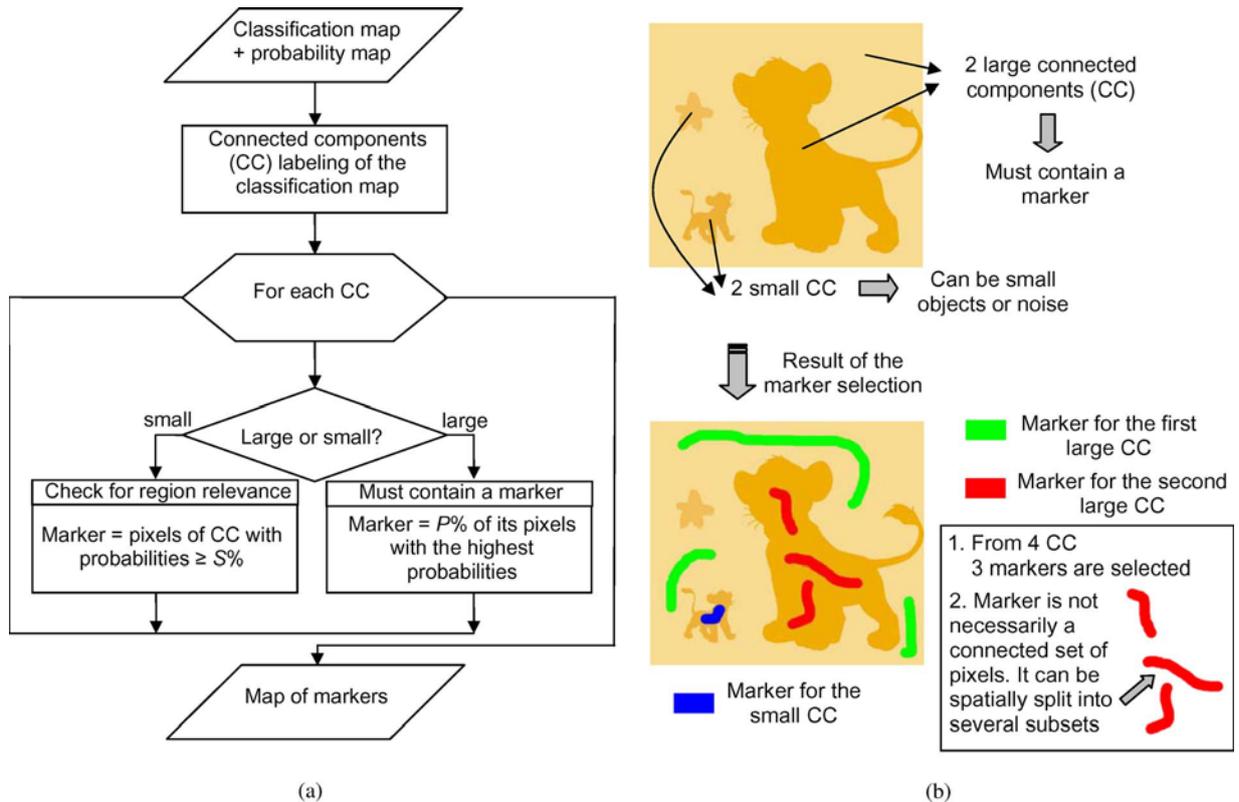


Fig. 3. (a) Flowchart of the proposed marker selection procedure. (b) Illustrative example of the marker selection.

the region (i.e., an area of the region) as a criterion of the region size. The threshold of the number of pixels defining if the region is large depends on the resolution of the image and typical sizes of the objects of interest. For instance, if the image of the volcano is considered (experimental results on the volcano image are illustrated in Section III), where the goal is to classify lavas of different eruption periods, it is known that the lava of different formations consists of large homogeneous regions. Therefore, it can be assumed that the regions representing structures (lavas) in the image scene have a size of at least 10 km^2 . Thus, for an airborne 20-m-resolution image, the threshold of $M = 20$ pixels for dividing the regions in the groups of large/small ones can be chosen.

2) A parameter P , defining the percentage of pixels within the large region to be used as markers, depends on the previous parameter. Since a marker for the large region

must be composed at least of one pixel, the following condition must be fulfilled: $P \geq 100\%/M$.

3) The last parameter S , which is a threshold of probability estimates defining potential markers for a small region, depends on the probability of the presence of small structures in the image (which also depends on the image resolution and the classes of interest) and the importance of the potential small structures (i.e., what is the cost of losing the small structures in the classification map). For instance, if we are interested in determining regions of different lava formations in the volcano image, the small objects in the image may have no importance for us, and a high value of S can be chosen. However, if the classification aims at determining regions of sick/damaged plants in the field, it may be important not to lose any small region of the damaged species. In this case, the threshold S must be relaxed.

In Section III-D, the dependence of the classification accuracies from the chosen parameters for the marker selection is investigated experimentally. As a conclusion, each connected set of pixels with the same class in the classification map provides either one or zero marker. One should stress that a marker is not necessarily a connected set of pixels: It can spatially be split into several subsets [see Fig. 3(b)].

C. Construction of an MSF

The previous two steps result in a map of markers defining regions of interest in the image. The next step consists in the grouping of all the image pixels into an MSF [22], where each tree is rooted on a classification-derived marker.

For this purpose, each pixel is considered as a vertex $v \in V$ of an undirected graph $G = (V, E, W)$, where V and E are the sets of vertices and edges, respectively, and W is a mapping of the set of edges E into \mathbb{R}^+ . Each edge $e_{i,j} \in E$ of this graph connects a couple of vertices i and j corresponding to the neighboring pixels (in the following, we simply call vertices as pixels). Furthermore, a weight $w_{i,j}$ is assigned to each edge $e_{i,j}$, which indicates the degree of dissimilarity between two pixels connected by this edge. Different dissimilarity measures can be used for computing weights of edges, such as vector norms, Spectral Angle Mapper (SAM), and spectral information divergence (SID) [36].

The $L1$ vector norm between two pixel vectors $\mathbf{x}_i = (x_{i1}, \dots, x_{iB})^T$ and $\mathbf{x}_j = (x_{j1}, \dots, x_{jB})^T$ is given as

$$L1(\mathbf{x}_i, \mathbf{x}_j) = \sum_{b=1}^B |x_{ib} - x_{jb}|. \quad (4)$$

The SAM distance between \mathbf{x}_i and \mathbf{x}_j determines the spectral similarity between two vectors by computing the angle between them. It is defined as

$$SAM(\mathbf{x}_i, \mathbf{x}_j) = \arccos \left(\frac{\sum_{b=1}^B x_{ib}x_{jb}}{\left[\sum_{b=1}^B x_{ib}^2 \right]^{1/2} \left[\sum_{b=1}^B x_{jb}^2 \right]^{1/2}} \right). \quad (5)$$

The SID measure [37] computes the discrepancy of probabilistic behaviors between the spectral signatures of two pixels. It is defined as

$$SID(\mathbf{x}_i, \mathbf{x}_j) = \sum_{b=1}^B \left\{ q_b(\mathbf{x}_i) \log \left[\frac{q_b(\mathbf{x}_i)}{q_b(\mathbf{x}_j)} \right] + q_b(\mathbf{x}_j) \log \left[\frac{q_b(\mathbf{x}_j)}{q_b(\mathbf{x}_i)} \right] \right\} \quad (6)$$

where

$$q_b(\mathbf{x}_i) = \frac{x_{ib}}{\sum_{l=1}^B x_{il}}. \quad (7)$$

Furthermore, more complex dissimilarity measures for image segmentation have been proposed in [11] and [38].

Given a connected graph $G = (V, E)$, a *spanning tree* $T = (V, E_T)$ of G is a connected graph without cycles such that

$E_T \subset E$. A *spanning forest* $F = (V, E_F)$ of G is a nonconnected graph without cycles such that $E_F \subset E$.

Given a graph $G = (V, E, W)$, the minimum spanning tree is defined as a spanning tree $T^* = (V, E_{T^*})$ of G such that the sum of the edge weights of T^* is minimal

$$T^* \in \arg \min_{T \in ST} \left\{ \sum_{e_{i,j} \in E_T} w_{i,j} \right\} \quad (8)$$

where ST is a set of all spanning trees of G .

Given a graph $G = (V, E, W)$, the *MSF* rooted on a set of m distinct vertices $\{t_1, \dots, t_m\}$ consists in finding a spanning forest $F^* = (V, E_{F^*})$ of G , such that each distinct tree of F^* is grown from one root t_i , and the sum of the edge weights of F^* is minimal [22]

$$F^* \in \arg \min_{F \in SF} \left\{ \sum_{e_{i,j} \in E_F} w_{i,j} \right\} \quad (9)$$

where SF is a set of all spanning forests of G rooted on $\{t_1, \dots, t_m\}$.

In order to obtain the MSF rooted on markers, m additional vertices t_i , $i = 1, \dots, m$, are introduced. Each extra vertex t_i is connected by the edge with a null weight to the pixels representing a marker i . Furthermore, an additional root vertex r is added and is connected by the null-weight edges to the vertices t_i . The minimum spanning tree of the constructed graph induces an MSF in G , where each tree is grown on a vertex t_i ; the MSF is obtained after removing the vertex r . An example of the construction of the MSF rooted on markers is shown in Fig. 4. Prim's algorithm can be used for building the MSF (see Algorithm 1) [39]. The efficient implementation of the algorithm using a binary min heap (for the implementation of a min-priority queue) is possible [40]; the resulting time complexity of the algorithm is $O(|E| \log |V|)$.

Algorithm 1 Prim's Algorithm

Require: Connected graph $G = (V, E, W)$

Ensure: Tree $T^* = (V^*, E^*, W^*)$

$V^* = \{v\}$, v is an arbitrary vertex from V

while $V^* \neq V$ **do**

 Choose edge $e_{i,j} \in E$ with minimal weight such that $i \in V^*$ and $j \notin V^*$

$V^* = V^* \cup \{j\}$

$E^* = E^* \cup \{e_{i,j}\}$

end while

Each tree in the MSF forms a region in the segmentation map (by mapping the resulting graph onto an image). Finally, a spectral-spatial classification map is obtained by assigning the class of each marker to all the pixels grown from this marker.

Thus, the proposed procedure of the construction of an MSF from region markers is a region growing method, which consists of the following steps: First, seed regions are chosen to belong to the segmentation and classification maps. Then, at each iteration, a new pixel i is added to the segmentation and classification maps so that the dissimilarity criterion between

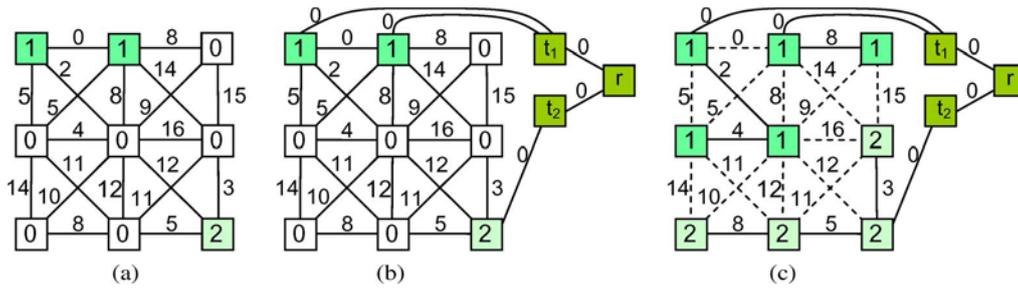


Fig. 4. Example of the construction of an MSF rooted on markers. (a) Original image graph G , where colored vertices represent markers 1 and 2; nonmarker pixels are denoted by “0.” (b) Addition of extra vertices t_1 , t_2 , and r to the graph. (c) Minimum spanning tree of the graph presented in (b); after removing the vertex r , an MSF is obtained, where each tree grown from the vertex t_i forms a region in the segmentation map.

this pixel and one of the pixels j already belonging to the segmentation/classification map is minimal. When including the new pixel to the classification map, a class of the pixel j is assigned to the pixel i .

D. Majority Voting Within Connected Components

Although the most reliable classified pixels are selected as markers, it may happen that a marker is classified to the wrong class. In this case, all the pixels within the region grown from this marker risk to be wrongly classified. In order to make the proposed classification scheme more robust, we propose to postprocess the classification map by applying a simple *majority voting* technique which has shown good performances for spectral-spatial classification [13], [14], [41]. For this purpose, connected-component labeling is applied on the obtained spectral-spatial classification map (using a four-neighborhood connectivity). Furthermore, for every connected component (region), all the pixels are assigned to the most frequent class when analyzing a pixelwise classification map within this region.

Note that an eight-neighborhood connectivity was used for the construction of an MSF, whereas a four-neighborhood connectivity was used for the majority voting. The use of the eight-neighborhood connectivity in the first case allows obtaining a more accurate (refined) segmentation map, without rough borders. Since an MSF is built from the set of markers, the number of regions does not depend on the chosen connectivity. When performing the last majority voting step, the use of the four-neighborhood connectivity results in a larger or the same number of connected components as the use of the eight-neighborhood connectivity. Therefore, the possible undersegmentation can be corrected in this step. One region from a segmentation map can be split into two connected regions when using the four-neighborhood connectivity. Furthermore, these two regions can be assigned to two different classes by the majority voting procedure.

III. EXPERIMENTAL RESULTS

Three different data sets were used for the experiments, with different contexts (agricultural, volcano, and urban areas) and acquired by different sensors (AVIRIS and ROSIS airborne imaging spectrometers). These data sets and the corresponding results are presented in the next three sections.

A. Classification of the Indiana Image

The *Indiana* image is of a vegetation area that was recorded by the AVIRIS sensor over the Indian Pine test site in Northwestern Indiana. The image has spatial dimensions of 145 by 145 pixels and a spatial resolution of 20 m/pixel. Twenty water absorption bands have been removed [42], and a 200-band image was used for the experiments. Sixteen classes of interest are considered, which are detailed in Table I, with a number of samples for each class in the reference data. A three-band false color image and the reference data are shown in Fig. 5. We have chosen randomly 50 samples for each class from the reference data as training samples, except for classes *alfalfa*, *grass/pasture-mowed*, and *oats*. These classes contain a small number of samples in the reference data. Therefore, only 15 samples for each of these classes were chosen randomly to be used as training samples. The remaining samples comprised the test set.

A pixelwise classification on the 200-band *Indiana* image was performed, using the multiclass one *versus* one SVM classifier with the Gaussian radial basis function (RBF) kernel. The optimal parameters C and γ were chosen by fivefold cross validation: $C = 128$ and $\gamma = 2^{-6}$. Global and class-specific accuracies are presented in Table I. Fig. 5 shows the obtained classification and probability maps.

Furthermore, a map of markers was created. For this purpose, labeling of connected components on the pixelwise classification map was performed, using the eight-neighborhood connectivity. For each connected component, the following are observed.

- 1) If it contained more than $M = 20$ pixels, $P = 5\%$ of its pixels with the highest probability estimates were selected as a marker for this component.
- 2) Otherwise, if a connected component contained pixels with the corresponding probability estimates not lower than the threshold S , these pixels were used as a marker.

In order to define a threshold S , the probability estimates for the whole image were sorted, and S was chosen equal to the lowest probability within the highest $T = 2\%$ of all probability estimates. The parameters for marker selection were chosen based on the following analysis: It is known that the image consists of the fields of different types of crops, i.e., large homogeneous regions. In the reference data, the class *oats* is represented by the smallest field of a size of 20 pixels. Therefore, the classification procedure must be able to

TABLE I
INFORMATION CLASSES, NUMBER OF LABELED SAMPLES (NO. OF SAMP.), AND CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE INDIANA IMAGE: OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), KAPPA COEFFICIENT (κ), AND CLASS-SPECIFIC ACCURACIES; "MV" MEANS INCLUDING A MAJORITY VOTING STEP

	No. of Samp.	SVM	Proposed Method (Using L1, SAM, SID Dissimilarity Measures)						M-WH	WH+MV
			L1	L1+MV	SAM	SAM+MV	SID	SID+MV		
OA	-	78.17	86.94	88.10	88.41	91.80	89.21	91.53	85.99	86.63
AA	-	85.97	88.66	91.82	91.57	94.28	92.08	94.14	86.95	91.61
κ	-	75.33	85.03	86.38	86.71	90.64	87.63	90.33	83.98	84.83
Corn-no till	1434	78.18	89.31	89.45	90.97	93.21	91.26	91.26	80.35	94.22
Corn-min till	834	69.64	66.71	62.63	69.52	96.56	74.62	95.79	71.94	78.06
Corn	234	91.85	64.13	99.46	95.65	95.65	95.65	95.65	73.37	88.59
Soybeans-no till	968	82.03	99.35	93.04	98.04	93.91	98.26	94.12	98.91	96.30
Soybeans-min till	2468	58.95	80.81	81.02	81.97	81.97	82.18	82.18	80.48	68.82
Soybeans-clean till	614	87.94	86.17	97.52	85.99	97.16	86.35	97.52	84.75	90.78
Alfalfa	54	74.36	92.31	92.31	94.87	94.87	94.87	94.87	94.87	94.87
Grass/pasture	497	92.17	93.51	93.51	94.63	94.63	94.63	94.63	95.30	95.08
Grass/trees	747	91.68	91.82	97.27	92.40	97.27	92.40	97.27	92.97	97.99
Grass/pasture-mowed	26	100	100	100	100	100	100	100	100	100
Hay-windrowed	489	97.72	99.77	99.77	99.77	99.77	99.77	99.77	99.54	99.54
Oats	20	100	100	100	100	100	100	100	100	100
Wheat	212	98.77	99.38	99.38	99.38	99.38	99.38	99.38	99.38	99.38
Woods	1294	93.01	98.47	98.47	97.59	99.68	99.52	99.52	99.36	97.11
Bldg-Grass-Tree-Drives	380	61.52	61.21	69.70	68.79	68.79	68.79	68.79	55.45	69.39
Stone-steel towers	95	97.78	95.56	95.56	95.56	95.56	95.56	95.56	64.44	95.56

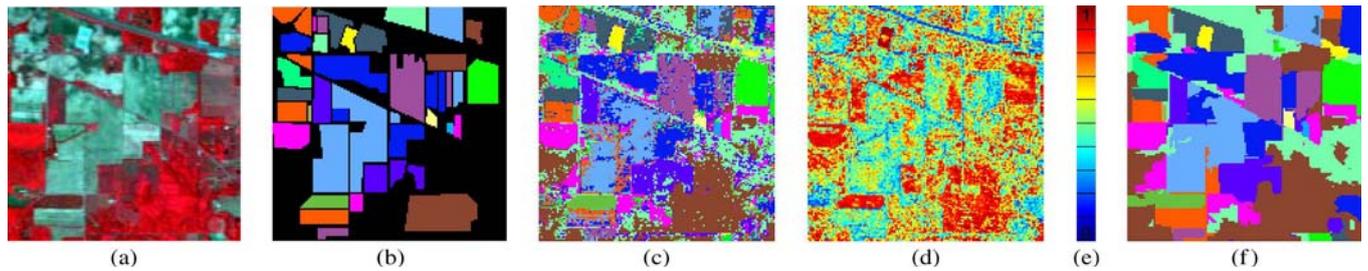


Fig. 5. Indiana image. (a) Three-band color composite (837, 636, and 537 nm). (b) Reference data: Corn-no till, Corn-min till, Corn, Soybeans-no till, Soybeans-min till, Soybeans-clean till, Alfalfa, Grass/pasture, Grass/trees, Grass/pasture-mowed, Hay-windrowed, Oats, Wheat, Woods, Bldg-Grass-Tree-Drives, Stone-steel towers. (c) Pixelwise classification map. (d) Probability map (probability estimates for each pixel to belong to the assigned class). (e) Scale of colors to represent the probability estimates in a probability map, from 0% probability at the bottom to 100% probability at the top. (f) Classification map obtained by the proposed scheme, using the SAM dissimilarity measure and including a majority voting step.

recognize the 20-pixel regions. We select $M = 20$. The parameter P is computed as $P = 100\%/M = 5\%$. This means that, for a region of 20 pixels, a one-pixel marker is selected. The last parameter T is chosen to be low since we know that the image does not contain small spatial structures. The correctness and robustness of this theoretical deduction of parameters are proved in Section III-D.

From 2250 connected components in the classification map, 107 markers were selected. Of the marker pixels presented in the reference data, 95% are correctly classified. In the next step, the image pixels were grouped into the MSF, built from the selected markers. We have investigated the use of different dissimilarity measures: the $L1$ vector norm,² the SAM, and the SID measures [36]. When the class of each marker was assigned to all the pixels of the corresponding tree, the spectral-spatial classification maps were obtained. Finally, the obtained classification maps were combined with the pixelwise classification

²We have also considered the $L2$ vector norm as a dissimilarity measure for the construction of the MSF. The corresponding classification accuracies are not given in this paper because of space limitations. These accuracies are mostly nonsignificantly lower than the ones obtained when using the $L1$ norm.

map, using the majority voting technique, as described in Section II-D.

Table I summarizes the accuracies of the pixelwise SVM and the proposed classification method (before and after the majority voting step). In order to compare performances of the proposed technique with the previously proposed methods, we have included results of a classification using majority vote within the adaptive neighborhoods defined by watershed segmentation ($WH + MV$) [13], as well as classification results obtained by performing watershed segmentation³ from the same set of markers ($M-WH$) [21]; these are recently proposed advanced techniques for spectral-spatial classification of hyperspectral images.

First of all, almost no oversegmentation is present in the obtained segmentation map (since one marker led to one region, a segmentation map contains 107 regions). As can be seen from Table I, both the global and most of the class-specific

³In [21], watershed segmentation is computed using the classical paradigm of the morphological image segmentation [19]: A gradient of the image is computed; then, a minima imposition technique is applied, followed by the watershed algorithm based on flooding simulations.

accuracies are improved by the proposed method. The majority voting step additionally improves most of the accuracies (except for the class *soybeans-no till*). The best global accuracies are achieved by the proposed method when using the SAM dissimilarity measure and including a majority voting step [the corresponding classification map is shown in Fig. 5(f)]. In this case, the overall accuracy is improved by 13.6 percentage points and the average accuracy by 8.3 percentage points when compared to the pixelwise classification. However, the use of other dissimilarity measures also leads to high accuracies; the highest accuracies for 7 of the 16 classes are achieved when using each of the proposed measures.

Furthermore, McNemar's test was performed in order to evaluate the statistical significance of differences in classification accuracies between the most accurate classification map ($SAM + MV$) and other classification maps [43]. According to the results of the test, the $SAM + MV$ classification accuracies are significantly different (here and in the following, we use 1% level of significance when reporting results of McNemar's test) from the accuracies of any other classification method applied for the *Indiana* image. The improvement of the accuracies after the majority voting step is also significant when using any of the applied dissimilarity measures.

B. Classification of the Hekla Image

The *Hekla* image was acquired by the AVIRIS sensor over the region surrounding the central volcano Hekla in Iceland [44]. The AVIRIS sensor operates in the wavelength range from 0.4 to 2.4 μm and utilizes four spectrometers collecting 224 data channels. During the data collection, spectrometer 4 was not working properly. The 64 data channels recorded by this spectrometer were deleted from the data, along with the first channels for the other three spectrometers (those channels were blank). Therefore, the 157 remaining data channels were used for the experiments. The considered image has spatial dimensions of 560 by 600 pixels. Twelve land cover classes of interest are considered, which are detailed in Table II, with a number of labeled samples for each class. Fig. 6 shows a three-band false color image and the reference data. Fifty samples for each class were randomly chosen from the reference data as training samples, and the rest of the samples were used as the test set.

A multiclass one *versus* one SVM classification on the original image was performed using the Gaussian RBF kernel. The parameters $C = 100$ and $\gamma = 0.1$ were determined by fivefold cross validation. Table II gives the classification accuracies, and the classification map is shown in Fig. 6(c).

In the next step, a map of markers was created, with the same parameters as for the *Indiana* image. Furthermore, segmentation and classification of the image were performed by constructing an MSF based on the selected markers. The obtained classification maps were further combined with the pixelwise classification map using the majority voting technique.

Table II gives the accuracies of the proposed classification method. As can be seen, both the global and most of the class-specific accuracies are improved when compared to the pixelwise classification. The majority voting step additionally improves the accuracies. As for the *Indiana* image, the best

global accuracies are achieved when performing the proposed classification with the SAM dissimilarity measure and the majority voting step. Fig. 6(d) shows the corresponding classification map, which is much less noisy than a pixelwise classification map. In this case, the overall accuracy is improved by 10.4 percentage points and the average accuracy by 9.0 percentage points when compared to the pixelwise classification. According to the results of McNemar's test, all the obtained classification maps are significantly different.

C. Classification of the University of Pavia Image

The proposed scheme was also tested on the *University of Pavia* image of an urban area, acquired by the ROSIS-03 optical sensor. The image is 610 by 340 pixels, with a spatial resolution of 1.3 m/pixel and 103 spectral channels. The reference data contain nine classes of interest. More information about the image, with the number of test and training samples for each class, can be found in [14].

Segmentation and classification of the *University of Pavia* image were performed using the proposed scheme. The parameters for an SVM classification were chosen by fivefold cross validation: $C = 128$ and $\gamma = 0.125$. Marker selection was performed with the same parameters as for the two previous data sets. Table III summarizes the classification accuracies for a pixelwise and spectral-spatial classification. In order to compare the performances of the proposed method with the previously proposed techniques, we have included in the table the accuracies of the classification of the *University of Pavia* image using an SVM, principal components, and extended morphological profiles; results are taken from [45]. This method has been recently proposed by Benediktsson *et al.* [46] and is considered as one of the most advanced methods for spectral-spatial classification of a multiband datum. Furthermore, the results of the spectral-spatial classification using majority voting within adaptive neighborhoods defined by spatial-based segmentation techniques are included. The following segmentation techniques are used for this purpose (leading to the best classification results among all the spatial-based methods): watershed segmentation ($WH + MV$) [13] and recursive hierarchical segmentation (RHSEG), with the possibility of merging only adjacent regions ($RHSEG0 + MV$) [9].

As can be seen from Table III, both the global and most of the class-specific accuracies are improved by the proposed method. The majority voting step additionally improves most of the accuracies. The best global accuracies are achieved by the proposed method when using the $L1$ vector norm for measuring dissimilarity between pixels. The corresponding classification map is significantly more accurate than any other obtained classification map, according to the results of McNemar's test. In this case, the overall accuracy is improved by 10.1 percentage points and the average accuracy by 6.5 percentage points when compared to the pixelwise classification. Those accuracies are higher than the ones obtained by the previously proposed techniques given for comparison. The use of the other two measures also led to the high classification accuracies for most of the classes. For instance, the use of the SAM measure led to the best accuracy of classification for the class *asphalt* when

TABLE II
INFORMATION CLASSES, NUMBER OF LABELED SAMPLES (NO. OF SAMP.), AND CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE HEKLA IMAGE: OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), KAPPA COEFFICIENT (κ), AND CLASS-SPECIFIC ACCURACIES; "MV" MEANS INCLUDING A MAJORITY VOTING STEP

	No. of Samp.	SVM	Proposed Method (Using L1, SAM, SID Dissimilarity Measures)					
			L1	L1+MV	SAM	SAM+MV	SID	SID+MV
OA	-	88.56	91.45	97.25	90.34	98.96	86.08	98.37
AA	-	89.44	94.11	96.36	94.89	98.45	91.44	98.25
κ	-	86.91	90.25	96.81	89.04	98.80	84.31	98.11
Andesite lava 1970	342	88.36	100	100	100	100	100	100
Andesite lava 1980 I	708	87.25	100	100	92.11	100	75.57	100
Andesite lava 1980 II	1496	88.24	98.96	98.96	96.96	99.86	97.10	100
Andesite lava 1991 I	2739	84.94	79.17	98.77	73.19	99.55	65.79	97.21
Andesite lava 1991 II	410	93.33	86.67	86.67	88.89	88.89	86.94	86.94
Andesite lava with moss cover	1023	94.24	98.46	98.46	98.46	98.46	98.46	98.46
Hyaloclastite formation	684	87.54	82.02	81.86	99.53	99.68	82.65	99.53
Lava covered with tephra and scoria	700	91.69	98.31	98.31	95.08	97.38	97.08	99.23
Rhyolite	404	85.88	92.37	96.05	96.89	100	96.05	100
Scoria	550	74.20	93.40	97.20	97.60	97.60	97.60	97.60
Firm and glacier ice	458	100	100	100	100	100	100	100
Snow	713	97.59	100	100	100	100	100	100

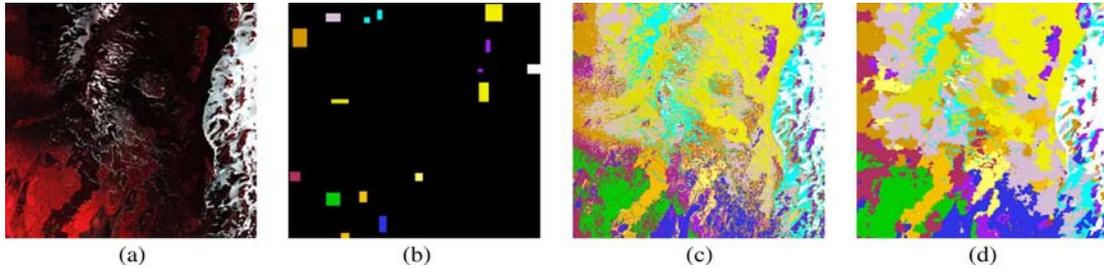


Fig. 6. Hekla image. (a) Three-band color composite (1125, 636, and 567 nm). (b) Reference data: **Andesite lava 1970**, **Andesite lava 1980 I**, **Andesite lava 1980 II**, **Andesite lava 1991 I**, **Andesite lava 1991 II**, **Andesite lava with moss cover**, **Hyaloclastite formation**, **Lava covered with tephra and scoria**, **Rhyolite**, **Scoria**, **Firm and glacier ice**, and snow (white). (c) Pixelwise classification map. (d) Classification map obtained by the proposed scheme, using the SAM dissimilarity measure and including a majority voting step.

TABLE III
CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE UNIVERSITY OF PAVIA IMAGE: OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), KAPPA COEFFICIENT (κ), AND CLASS-SPECIFIC ACCURACIES; "MV" MEANS INCLUDING A MAJORITY VOTING STEP

	SVM	Proposed Method (Using L1, SAM, SID Dissimilarity Measures)						WH +MV	RHSEG0 +MV	EMP
		L1	L1+MV	SAM	SAM+MV	SID	SID+MV			
OA	81.01	84.14	91.08	78.62	83.94	77.97	83.38	85.42	90.00	85.22
AA	88.25	92.35	94.76	85.90	88.01	85.94	89.27	91.31	94.15	90.76
κ	75.86	79.71	88.30	73.28	79.59	72.45	78.94	81.30	86.86	80.86
Asphalt	84.93	93.05	93.16	95.78	96.88	95.69	96.13	93.64	73.33	95.36
Meadows	70.79	72.30	85.65	62.61	70.92	61.60	68.63	75.09	88.73	80.33
Gravel	67.16	89.15	89.15	94.66	77.69	94.27	95.21	66.12	97.47	87.61
Trees	97.77	87.02	91.24	96.26	96.50	93.30	93.75	98.56	98.45	98.37
Metal sheets	99.46	99.91	99.91	88.05	99.82	99.82	99.82	99.91	99.10	99.48
Bare soil	92.83	97.11	99.91	85.13	99.80	83.53	99.76	97.35	98.43	63.72
Bitumen	90.42	98.57	98.57	97.55	97.55	97.96	97.96	96.23	95.92	98.87
Bricks	92.78	95.66	99.05	96.70	98.63	96.88	98.75	97.92	98.81	95.41
Shadows	98.11	98.36	96.23	56.35	54.34	50.44	53.46	96.98	97.11	97.68

compared to all the previous results found in the literature. However, the global accuracies are not as high as when using the $L1$ vector norm mainly because of the low accuracy for the class *shadows*. The SAM distance is actually designed with the purpose that the poorly illuminated and more brightly illuminated pixels from the same class would be mapped to the same spectral angle despite the difference in illumination. It can be an explanation of the fact why the SAM distance led to the assimilation of the shadowed regions with the neighboring structures.

D. Assessment of the Robustness of the Parameter Settings

In Section II-B, the procedure of the choice of parameters for the marker selection procedure was explained. In this section, the dependence of the classification accuracies from the chosen parameters is investigated experimentally for the *Indiana* and the *Hekla* images.

First, the choice of the parameter P is analyzed. For this purpose, the other two parameters were fixed as $M = 20$ and $T = 2$. Classification using the proposed scheme was performed, with the parameter P varying from 5 to 100. The $L1$ vector

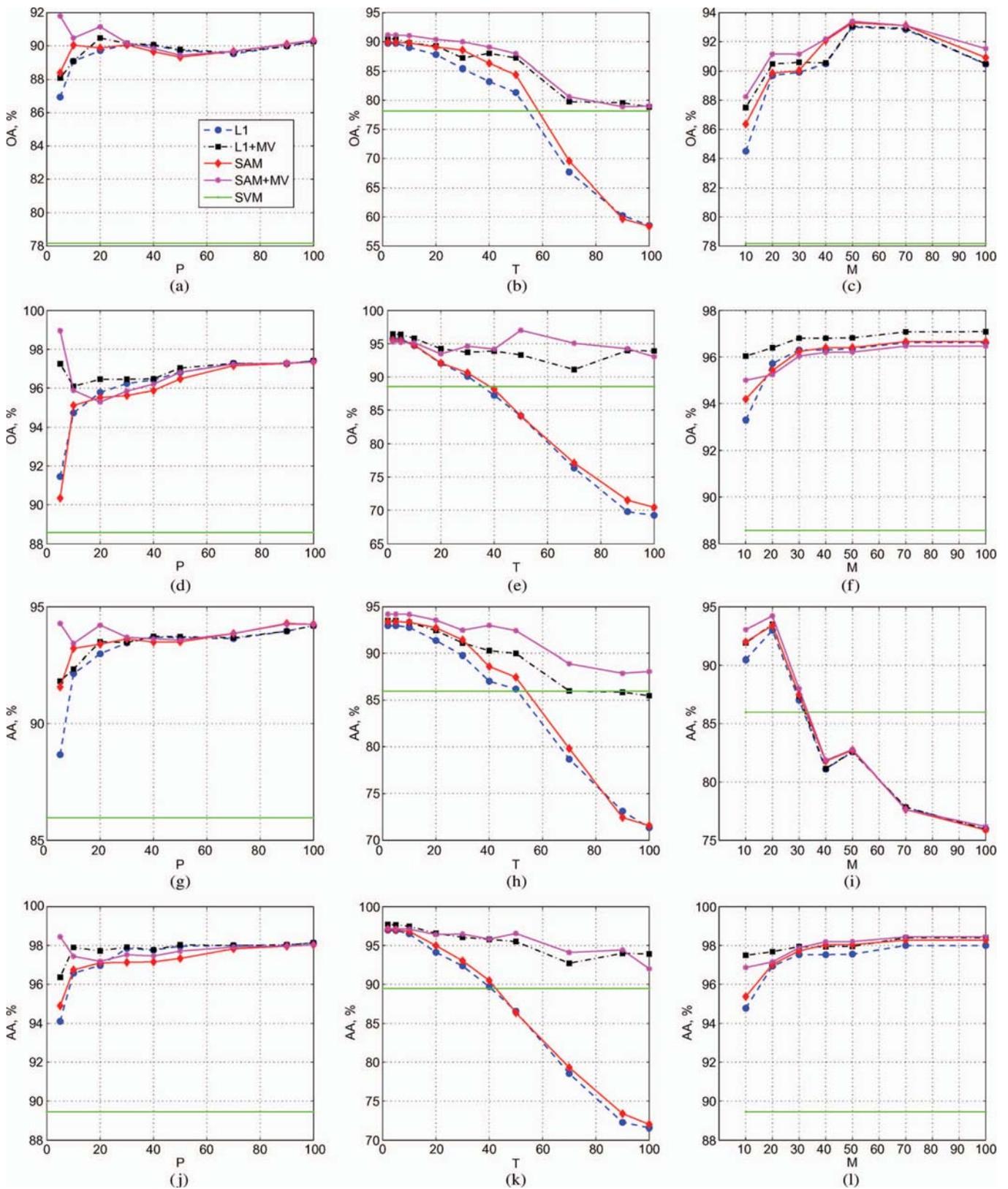


Fig. 7. Assessment of the robustness of the parameter settings. (a)–(f) Overall and (g)–(l) average classification accuracies as functions of parameters for the marker selection procedure [(a), (d), (g), (j)] P ($M = 20, T = 2$), [(b), (e), (h), (k)] T ($M = 20, P = 20$), and [(c), (f), (i), (l)] M ($P = 20, T = 5$), for the [(a)–(c), (g)–(i)] Indiana and [(d)–(f), (j)–(l)] Hekla images.

norm and the SAM dissimilarity measures were used. Fig. 7 shows the obtained overall [Fig. 7(a) and (d)] and average [Fig. 7(g) and (j)] classification accuracies for both data sets. As

can be seen from the figures, for any value of P , the accuracies are significantly improved when compared to the pixelwise classification. When the value of P is low (a few marker pixels

are chosen), the construction of the MSF from the markers gives less accurate classification maps than when it is high. However, the postprocessing majority voting step improves the accuracies significantly. Thus, the complete classification procedure gives the best performances for a low value of P . Then, with the increase of P , the classification accuracies converge to a constant value, for both dissimilarity measures, and the majority step does not additionally increase the accuracies. Results are coherent for both images. As a conclusion, this parameter has no significant influence on the classification accuracies. When a few marker pixels are chosen (for low values of P), the construction of an MSF using the SAM distance gives better segmentation results (provides more accurate regions) than a pixelwise classification. Therefore, after the additional processing of the corresponding classification maps by majority voting, the best classification results are achieved.

The choice of the parameter T is investigated, by fixing parameters $M = 20$ and $P = 2$, varying T from 2 to 100, and performing classification for both data sets. The overall and average classification accuracies are shown in Fig. 7 (second column). As can be seen, the results are coherent for both dissimilarity measures and both data sets. The highest classification accuracies are achieved for the low values of T . These results are coherent with the theoretical analysis given in the previous sections. Since the *Indiana* and the *Hekla* images contain large spatial structures (of crop fields and lava formations, respectively), markers for small regions must be selected only in the case if the probability of their correct classification is very high. Therefore, with the increase of T , the probability of selecting a marker for a misclassified region increases, and the classification accuracies decrease. The majority voting step becomes more important for higher values of T , and for the *Hekla* image, the accuracies for the complete proposed procedure (including the majority voting step) remain almost unchanged, independently of the parameter T .

Furthermore, the choice of the parameter M is investigated, with the other parameters being fixed ($P = 20$ and $T = 5$). Fig. 7 (third column) shows the overall and average accuracies of classification when varying the parameter M from 10 to 100. For the *Hekla* image, the accuracies are slightly lower for $M = 10$, when compared to other values of this parameter, for which the accuracies remain almost unchanged. This is due to the fact that the image contains mostly large regions of lava fields. Therefore, for any value of M higher than 20, the classification procedure shows a good performance. For the *Indiana* image, with the increase of the parameter M up to 50, the overall accuracy slightly increases. However, the average accuracy reaches its maximum for $M = 20$ and then significantly decreases. These results confirm the theoretical analysis: The smallest crop field in the image scene has a size of 20 pixels. When a high value of M for identifying significant regions is chosen, small crop fields risk to be assimilated with the neighboring structures. If this happens, majority voting cannot reconstitute these regions, and they disappear from the final classification map.

In conclusion, the experimental analysis of the parameter choice for the marker selection procedure has confirmed that, by using some *a priori* information for the image, pa-

rameters leading to good classification performances can be deducted.

IV. CONCLUSION

A large number of spectral channels in a hyperspectral image increase the potential of discriminating physical materials and structures in a scene. However, it presents challenges to image analysis because of the huge volume of data that the hyperspectral image usually consists of. Although pixelwise classification techniques have given high classification accuracies when dealing with hyperspectral data, the incorporation of the spatial information into the classifier is needed for further improvement of the classification accuracies.

In this paper, a new spectral-spatial classification scheme for hyperspectral images has been proposed. The method is based on the construction of an MSF, rooted on the markers selected by using pixelwise classification results. Experimental results, presented on the three data sets, have shown that the proposed method improves the classification accuracies, when compared to previously proposed classification schemes, and provides accurate segmentation and classification maps.

Different distances have been investigated for measuring the dissimilarity between pixels when constructing an MSF. It is shown that, in most of the cases, the SAM distance gives the best performances. However, this distance does not discriminate some particular classes from urban areas, such as shadows, and assimilate the pixels belonging to these classes with the neighboring structures.

In conclusion, the proposed classification methodology succeeded in taking advantage of the spatial and the spectral information simultaneously. The method performs well for images representing different scenes: those containing large spatial structures with spectrally confusing classes and those containing small and complex structures. Furthermore, its efficient implementation is possible.

ACKNOWLEDGMENT

The authors would like to thank P. Gamba and D. Landgrebe for providing the hyperspectral data.

REFERENCES

- [1] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. New York: McGraw-Hill, 1995.
- [2] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*. New York: Wiley-Interscience, 2007.
- [3] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. New York: Wiley, 2003.
- [4] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agric.*, vol. 39, no. 2, pp. 67–93, 2003.
- [5] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [6] M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data," Ph.D. dissertation, Grenoble Inst. Technol., Grenoble, France, 2007.
- [7] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

- [8] A. Farag, R. Mohamed, and A. El-Baz, "A unified framework for map estimation in remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, Jul. 2005.
- [9] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Classification of hyperspectral data using support vector machines and adaptive neighborhoods," in *Proc. 6th EARSeL SIG IS Workshop*, Tel Aviv, Israel, 2009.
- [10] M. Zhang, L. O. Hall, and D. B. Goldgof, "A generic knowledge-guided image segmentation and labeling system using fuzzy clustering algorithms," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 32, no. 5, pp. 571–582, Oct. 2002.
- [11] S. Makrogiannis, G. Economou, and S. Fotopoulos, "A region dissimilarity relation that combines feature-space and spatial information for color image segmentation," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 1, pp. 44–53, Feb. 2005.
- [12] R. Gonzalez and R. Woods, *Digital Image Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [13] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in *Proc. IGARSS*, Boston, MA, 2008, pp. III-652–III-655.
- [14] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.
- [15] P. Soille, *Morphological Image Analysis*, 2nd ed. New York: Springer-Verlag, 2003.
- [16] A. C. Jalba, M. H. Wilkinson, and J. B. T. M. Roerdink, "Automatic segmentation of diatom images for classification," *Microsc. Res. Tech.*, vol. 65, no. 1/2, pp. 72–85, Sep. 2004.
- [17] G. Noyel, J. Angulo, and D. Jeulin, "Morphological segmentation of hyperspectral images," *Image Anal. Stereol.*, vol. 26, pp. 101–109, 2007.
- [18] O. Gómez, J. A. González, and E. F. Morales, "Image segmentation using automatic seeded region growing and instance-based learning," in *Proc. 12th Iberoamerican Congr. Pattern Recog.*, Valparaiso, Chile, Nov. 2007, pp. 192–201.
- [19] G. Noyel, "Filtrage, Réduction de Dimension, Classification et Segmentation Morphologique Hyperspectrale," Ph.D. dissertation, Center Math. Morphology, Paris School Mines, Fontainebleau, France, 2008.
- [20] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data. An Introduction to Cluster Analysis*. Hoboken, NJ: Wiley, 1990.
- [21] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Classification based marker selection for watershed transform of hyperspectral images," in *Proc. IGARSS*, Cape Town, South Africa, 2009.
- [22] J. Stawiaski, "Mathematical morphology and graphs: Application to interactive medical image segmentation," Ph.D. dissertation, Paris School Mines, Paris, France, 2008.
- [23] O. J. Morris, M. J. Lee, and A. G. Constantinides, "Graph theory for image analysis: An approach based on the shortest spanning tree," *Proc. Inst. Elect. Eng.—Commun., Radar Signal Process.*, vol. 133, no. 2, pp. 146–152, Apr. 1986.
- [24] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [25] W. Tao, H. Jin, and Y. Zhang, "Color image segmentation based on mean shift and normalized cuts," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 5, pp. 1382–1389, Oct. 2007.
- [26] C. Allène, J. Audibert, M. Couprie, J. Cousty, and R. Keriven, "Some links between min-cuts, optimal spanning forests and watersheds," in *Proc. 8th Int. Symp. Math. Morphol.*, Rio de Janeiro, Brazil, Oct. 2007, pp. 253–264.
- [27] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest," in *Proc. ICCV*, Kyoto, Japan, 2009, pp. 731–738.
- [28] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
- [29] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Evaluation of kernels for multiclass classification of hyperspectral remote sensing data," in *Proc. ICASSP*, May 2006, pp. II-813–II-816.
- [30] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba, "Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRS-S data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3857–3865, Nov. 2009.
- [31] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *J. Mach. Learn. Res.*, vol. 5, pp. 975–1005, Dec. 2004.
- [32] C. Chang and C. Lin, LIBSVM: A Library for Support Vector Machines, 2001. [Online]. Available: <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [33] H.-T. Lin, C.-J. Lin, and R. C. Weng, "A note on Platt's probabilistic outputs for support vector machines," Dept. Comput. Sci., Nat. Taiwan Univ., Taipei, Taiwan, 2003.
- [34] J. Platt, "Probabilistic outputs for support vector machines and comparison to regularized likelihood methods," in *Advances in Large Margin Classifiers*, A. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans, Eds. Cambridge, MA: MIT Press, 2000.
- [35] L. Shapiro and G. Stockman, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 2002.
- [36] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *Int. J. Appl. Earth Observation Geoinformation*, vol. 8, no. 1, pp. 3–17, 2006.
- [37] C.-I. Chang, "An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis," *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1927–1932, Aug. 2000.
- [38] P. Corsini, B. Lazzerini, and F. Marcelloni, "A fuzzy relational clustering algorithm based on a dissimilarity measure extracted from data," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 1, pp. 775–781, Feb. 2004.
- [39] R. C. Prim, "Shortest connection networks and some generalizations," *Bell Syst. Tech. J.*, vol. 36, pp. 1389–1401, 1957.
- [40] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. Cambridge, MA: MIT Press, 2001.
- [41] A. Widayati, B. Verbist, and A. Meijerink, "Application of combined pixel-based and spatial-based approaches for improved mixed vegetation classification using IKONOS," in *Proc. 23rd Asian Conf. Remote Sens.*, 2002, p. 8.
- [42] S. Tadjudin and D. A. Landgrebe, "Covariance estimation with limited training samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 4, pp. 2113–2118, Jul. 1999.
- [43] G. M. Foody, "Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 5, pp. 627–633, 2004.
- [44] J. A. Benediktsson and I. Kanellopoulos, "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1367–1377, May 1999.
- [45] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009, Supplement 1.
- [46] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.



Yuliya Tarabalka (S'08) received the B.S. degree in computer science from Ternopil State Ivan Pul'uj Technical University, Ternopil, Ukraine, in 2005 and the M.Sc. degree in signal and image processing from the Grenoble Institute of Technology (INPG), Grenoble, France, in 2007. She is currently working toward the Ph.D. degree at the University of Iceland, Reykjavik, Iceland, and at INPG.

From July 2007 to January 2008, she was a Researcher with the Norwegian Defence Research Establishment, Norway. Her current research work is funded by the "Hyper-I-Net" Marie Curie Research Training Network. Her research interests are in the areas of image processing, pattern recognition, hyperspectral imaging, and development of efficient algorithms.



Jocelyn Chanussot (SM'04) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995 and the Ph.D. degree from the University of Savoie, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory, Delegation Generale de l'Armement (French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, was an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA Lab), INPG. His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was the Cochair of the GRS Data Fusion Technical Committee (2005–2008) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008). He was the Founding President of the IEEE GRS French Chapter (2007).



Jón Atli Benediktsson (F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984 and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively.

He is currently the Pro Rector for Academic Affairs and a Professor of electrical and computer engineering with the University of Iceland. His research interests are in remote sensing, pattern recognition, neural networks, image processing, and signal processing, and he has published extensively in those

fields.

Dr. Benediktsson is a member of Societas Scinetiarum Islandica and Tau Beta Pi. He was an Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2003 to 2008. He has been on the AdCom of the IEEE Geoscience and Remote Sensing Society (GRSS) since 2000 and is the GRSS Executive Vice President for 2010–2011. He was the recipient of the Stevan J. Kristof Award from Purdue University in 1991 as an outstanding graduate student in remote sensing, the Icelandic Research Council's Outstanding Young Researcher Award in 1997, the IEEE Third Millennium Medal in 2000, the yearly research award from the Engineering Research Institute of the University of Iceland in 2006, and the Outstanding Service Award from the IEEE GRSS in 2007. He was also a corecipient of the University of Iceland's Technology Innovation Award in 2004.

Paper 6

Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton,
“Multiple spectral-spatial classification approach for hyperspectral data,”
IEEE Trans. on Geoscience and Remote Sensing, under review.

Multiple Spectral-Spatial Classification Approach for Hyperspectral Data

Yuliya Tarabalka^{1,2} *Student Member, IEEE*, Jón Atli Benediktsson¹ *Fellow, IEEE*,
Jocelyn Chanussot² *Senior Member, IEEE*, and James C. Tilton³ *Senior Member, IEEE*,

¹Faculty of Electrical and Computer Engineering – University of Iceland
Hjardarhaga 2-6, 107 Reykjavik, Iceland

²GIPSA-Lab – Grenoble Institute of Technology,
Domaine Universitaire – BP 46 – 38402 Saint-Martin-d’Hères Cedex, France

³NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

e-mail: yuliya.tarabalka@hyperinet.eu, benedikt@hi.is, jocelyn.chanussot@gipsa-lab.grenoble-inp.fr,
james.c.tilton@nasa.gov

Abstract—A new multiple classifier approach for spectral-spatial classification of hyperspectral images is proposed. Several classifiers are used independently to classify an image. For every pixel, if all the classifiers have assigned this pixel to the same class, the pixel is kept as a marker, i.e., a seed of the spatial region, with the corresponding class label. We propose to use spectral-spatial classifiers at the preliminary step of the marker selection procedure, each of them combining the results of a pixelwise classification and a segmentation map. Different segmentation methods based on dissimilar principles lead to different classification results. Furthermore, a minimum spanning forest is built, where each tree is rooted on a classification-driven marker and forms a region in the spectral-spatial classification map. Experimental results are presented for two hyperspectral airborne images. The proposed method significantly improves classification accuracies, when compared to previously proposed classification techniques.

Index Terms—Hyperspectral images, classification, segmentation, multiple classifiers, minimum spanning forest

I. INTRODUCTION

Hyperspectral imaging is a relatively recent technique in remote sensing. Acquired remotely by airborne or spaceborne sensors, hyperspectral data are comprised of hundreds of spatially co-registered images corresponding to different spectral channels [1], [2]. Figure 1 illustrates the structure of a hyperspectral image. Every pixel is presented as a B -dimensional feature vector across the wavelength dimension, called the spectrum of the material in this pixel. This rich information in every spatial location increases the capability to distinguish different physical materials. Thus, hyperspectral imagery opens new perspectives for image classification, which is an important task for a wide variety of applications (precision agriculture, monitoring and management of the environment, security issues).

However, such a large number of spectral channels, usually coupled with limited availability of reference data ¹, presents

This research is supported in part by the Marie Curie Research Training Network “HYPER-I-NET”.

¹By *reference data*, we mean manually labeled pixels which are used for training classifiers followed by assessment of classification accuracies.

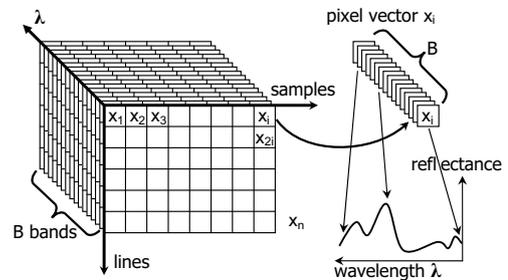


Fig. 1. Structure of a hyperspectral image.

challenges to image analysis. While pixelwise classification techniques process each pixel independently without considering information about spatial structures [3], [4], [5], further improvement of classification results can be achieved by considering spatial dependencies between pixels, i.e., by performing spectral-spatial classification [6], [7], [8], [9], [10].

Segmentation techniques, partitioning an image into homogeneous regions with respect to some criterion of interest (called homogeneity criterion, e.g., intensity or texture), are powerful tools for defining spatial dependencies [11]. In previous works, we have distinguished spatial structures in the hyperspectral image by performing unsupervised segmentation [9], [12], [13]. Watershed, partitional clustering and Hierarchical SEGmentation (HSEG) techniques have been used for this purpose. Segmentation and pixelwise classification were applied independently, then results were combined using a majority voting rule (see Figure 2). Thus, every region from a segmentation map was considered as an adaptive homogeneous neighborhood for all the pixels within this region. The described technique led to a significant improvement of classification accuracies and provided more homogeneous classification maps, when compared to classification techniques using local neighborhoods in order to include spatial information into a classifier.

The drawback of unsupervised image segmentation is that

the segmentation results depend on the chosen measure of homogeneity which is image-dependent [11]. Too relaxed or too restricted homogeneity criterion can lead to undersegmented (several regions are detected as one) or oversegmented (one region is detected as several ones) results, respectively.

An alternative way to get accurate segmentation results consists in performing a marker-controlled segmentation [11], [14]. The idea behind this approach is to select for every spatial object one or several pixels belonging to this object (called a *region seed*, or a *marker* of the corresponding region) and to grow regions from the selected seeds, so that every region in the resulting segmentation map is associated with one region seed. The markers of regions can be chosen either manually, or automatically. Recently we have proposed to use probability estimates obtained by the pixelwise Support Vector Machines (SVM) classification in order to select the most reliable classified pixels as markers, i.e., seeds of spatial regions [15]. Furthermore, image pixels were grouped into a Minimum Spanning Forest (MSF), where each tree was rooted on a classification-derived marker. The decision to connect a pixel, which was not yet in the forest, to one of the trees in the forest was based on its similarity to one of the adjacent pixels already belonging to the forest. By assigning the class of each marker to all the pixels within the region grown from this marker, a spectral-spatial classification map was obtained. The described technique led to a significant improvement of classification accuracies when compared to previously proposed methods. The drawback of this method is that the choice of markers strongly depends on the performances of the selected pixelwise classifier (e.g., the SVM classifier in our previous work [15]).

In this work, we aim to mitigate the dependence of the marker selection procedure from the choice of a pixelwise classifier. This can be achieved by using not a single classification algorithm for marker selection, but an ensemble of classifiers, i.e., *multiple classifiers*. For this purpose, several individual classifiers must be chosen and combined within one system in such a way that the complementary benefits of each classifier are used, while their weaknesses are avoided.

In this paper, a new *marker selection method* based on a *multiple classifier (MC)* system is proposed. Several classifiers are used independently to classify an image. Furthermore, a marker map is constructed by selecting the pixels assigned by all the classifiers to the same class. We propose to use spectral-spatial classifiers in the preliminary step of the marker selection procedure, each of them combining the results of a pixelwise classification and one of the unsupervised segmentation techniques (see Figure 2). By using spectral-spatial classifiers in this step, spatial context in the image is taken into account, and classification maps are more accurate when compared to pixelwise classification maps. This leads to more accurate marker selection results. The proposed marker selection method is incorporated into a new *Multiple Spectral-Spatial Classification scheme (MSSC-MSF)* based on the construction of an MSF from region markers.

In order to assess the importance of spectral-spatial approaches for marker selection, we have also implemented a *Multiple Classification scheme (MC-MSF)*. Here, spectral-spatial classification maps are replaced by the maps obtained

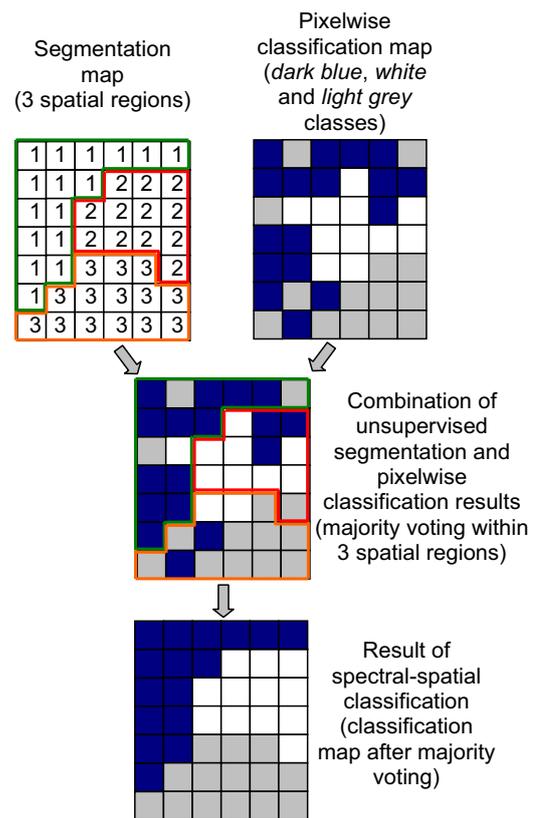


Fig. 2. Example of spectral-spatial classification using majority voting within segmentation regions.

using pixelwise classification techniques. Finally, a marker map is computed and an MSF from the selected markers is constructed.

Although the classification approach proposed in this paper has been designed for hyperspectral data, the method is general and can be applied for other types of data as well. Two hyperspectral airborne images are used to demonstrate experimental results: an image recorded by the Reflective Optics System Imaging Spectrometer (ROSIS) over the University of Pavia, Italy, and an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) image acquired over Northwestern Indiana's Indian Pines site [16].

The paper is organized as follows. In the next section, the multiple classifier approach is briefly discussed. Section III describes the proposed classification scheme. Experimental results are discussed in Section IV. Finally, conclusions are drawn in Section V.

II. MULTIPLE CLASSIFIER APPROACH

The traditional approach for a pattern recognition problem is to search for the individual algorithm giving the best possible classification performances. However, in many cases, the classification accuracy can be improved by using an ensemble of classifiers, or multiple classifiers. This is due to the fact that although one of the classification algorithms

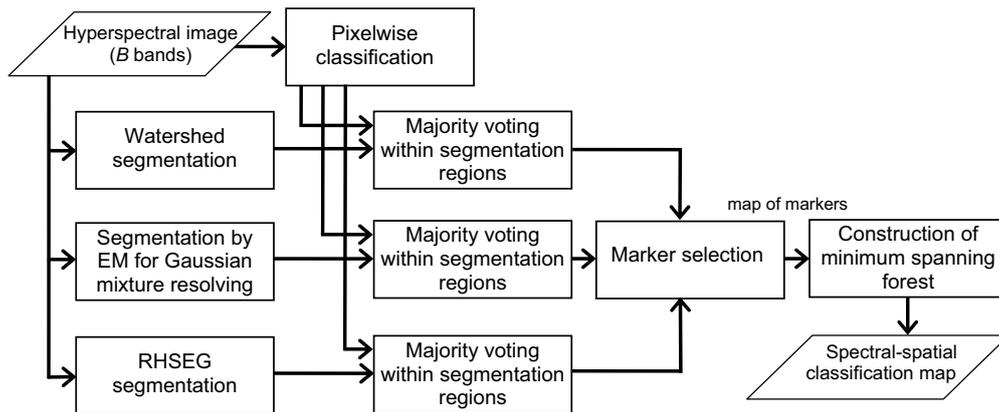


Fig. 3. Flow-chart of the proposed *MSSC-MSF* classification scheme.

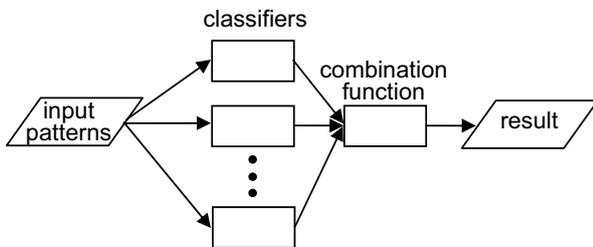


Fig. 4. Schematic diagram of a multiple classifier system.

would yield the best performances, the sets of pixels (patterns in general) misclassified by the different algorithms would not necessarily overlap. Thus, the aim of an *MC* system is to determine an efficient combination method that makes use of the complementary benefits of each classifier, while tackling the individual drawbacks [17], [18], [19].

A schematic representation of an *MC* system is given in Figure 4. An important issue for an efficient *MC* system is that the individual classifiers should be independent. More precisely, the classifiers should not agree with each other when they misclassify a pixel [18]. The complementary properties of the different classifiers selected for the *MC* system should ensure to a certain extent this requirement.

Another important issue is the rule for combining the individual classifiers (i.e., combination function). The individual classifier outputs, such as class labels and possibly posterior probabilities, are typically combined by voting rules, belief functions, statistical techniques, the Dempster-Shafer evidence theory, and other schemes [17]. For a given pixel, if all the classifiers agree on the same class k , the evident combination rule consists in assigning this pixel to the class k in the final classification map. On the other side, when individual classifiers disagree in assigning the given pixel, the procedure of final decision making is not that straightforward, and different combination functions may yield different results. A

typical result of the *MC* system is a final classification map, where each pixel has a unique class label. This type of *MS* systems has been previously used for remote sensing image classification [19], [20], [21].

In this paper, we propose to address the combination rule issue in the following way: According to the exclusionary rule, only the pixels where all the classifiers agree, i.e., the most reliable pixels, are kept in the classification map. The rest of the pixels are further classified by constructing an *MSF* rooted on the “reliable” pixels, i.e., by incorporating the spatial information into classification.

Coming back to the first issue for designing an *MC* system, different individual classifiers must be chosen. For instance, standard pixelwise classification algorithms can be used for this purpose, such as SVM, Maximum Likelihood (*ML*), k -Nearest Neighborhood (k -*NN*) methods (parametric and non-parametric techniques, based on different principles). We have used these individual techniques in the *MC-MSF* classification scheme. Furthermore, we propose to use spectral-spatial classifiers as individual classifiers for the *MC* system (*MSSC-MSF* classification scheme), each of them combining the results of a pixelwise classification and one of the unsupervised segmentation techniques. Different segmentation methods based on dissimilar principles lead to different classification results. The use of spectral-spatial classifiers yields more accurate classification maps, when compared to those obtained by performing pixelwise classification.

III. PROPOSED CLASSIFICATION SCHEME

The flow-chart of the proposed *MSSC-MSF* classification method is depicted in Figure 3. At the input a B -band hyperspectral image is given, which can be considered as a set of n pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbb{R}^B, j = 1, 2, \dots, n\}$. Classification consists in assigning each pixel to one of the K classes of interest. In the following, each step of the proposed procedure is described.

Segmentation can be defined as an exhaustive partitioning of the input image into regions, each of which is considered to

be homogeneous with respect to some criterion of interest. We have investigated the use of three techniques for hyperspectral image segmentation, as described hereafter.

A. Watershed segmentation

Watershed transformation is a powerful morphological approach to image segmentation which combines region growing and edge detection. The watershed is usually applied to the gradient function, and it divides an image into regions, so that each region is associated with one minimum of the gradient image [22].

The extension of a watershed technique to the case of hyperspectral images has been investigated in [13], [23]. In this paper, we present watershed results, obtained by the scheme we proposed and described in [13]: First, a one-band Robust Color Morphological Gradient (RCMG) [24] for the hyperspectral image is computed.

For each pixel vector \mathbf{x}_p , let $\chi = [\mathbf{x}_p^1, \mathbf{x}_p^2, \dots, \mathbf{x}_p^e]$ be a set of e vectors contained within a structuring element E (i.e., the pixel \mathbf{x}_p itself and $e - 1$ neighboring pixels). A 3×3 square structuring element with the origin in its center is typically used. The *Color Morphological Gradient (CMG)*, using the Euclidean distance, is computed as:

$$CMG_E(\mathbf{x}_p) = \max_{i,j \in \chi} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}, \quad (1)$$

i.e., the maximum of the distances between all pairs of vectors in the set χ . One of the drawbacks of the CMG is that it is very sensitive to noise. To overcome the problem of outliers, the RCMG has been proposed [24]. The scheme to make a CMG robust consists of removing the two pixels that are furthest apart and then finding the CMG of the remaining pixels. This process can be repeated several times until a good estimate of the gradient is obtained.

Thus, the RCMG, using the Euclidean distance, can be defined as:

$$RCMG_E(\mathbf{x}_p) = \max_{i,j \in [\chi - REM_r]} \{\|\mathbf{x}_p^i - \mathbf{x}_p^j\|_2\}, \quad (2)$$

where REM_r is a set of r vector pairs removed. If E is a 3×3 square structuring element, $r = 1$ is recommended [24].

Furthermore, watershed transformation is applied on the gradient image, using a standard algorithm [25]. As a result, the image is partitioned into a set of regions, and one subset of watershed pixels, i.e., pixels situated on the borders between regions. Finally, every watershed pixel is assigned to the neighboring region with the ‘‘closest’’ median² (the distance between the vector median of this region and the watershed pixel is minimal).

B. Segmentation by expectation maximization

The Expectation Maximization (EM) algorithm for the Gaussian mixture resolving belongs to the group of partitional clustering techniques [12], [27]. The use of partitional clustering for hyperspectral image segmentation has been discussed

²A standard vector median [26] for the region $\mathbf{S} = \{\mathbf{s}_j \in \mathbb{R}^B, j = 1, 2, \dots, l\}$ is defined as $\mathbf{s}_{VM} = \arg \min_{\mathbf{s} \in \mathbf{S}} \{\sum_{j=1}^l \|\mathbf{s} - \mathbf{s}_j\|_1\}$

in [12]. Clustering aims at finding groups of spectrally similar pixels. We assume that pixels belonging to the same cluster are drawn from a multivariate Gaussian probability distribution. Each image pixel can be statistically modelled by the following probability density function:

$$p(\mathbf{x}) = \sum_{c=1}^C \omega_c \phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \quad (3)$$

where C is the number of clusters, $\omega_c \in [0, 1]$ is the mixing proportion (weight) of cluster c with $\sum_{c=1}^C \omega_c = 1$, and $\phi(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate Gaussian density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$\phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) = \frac{1}{(2\pi)^{B/2}} \frac{1}{|\boldsymbol{\Sigma}_c|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c^{-1} (\mathbf{x} - \boldsymbol{\mu}_c)\}. \quad (4)$$

The parameters of the distributions $\psi = \{C, \omega_c, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c; c = 1, 2, \dots, C\}$ are estimated by the EM algorithm, as described in [12]. An upper bound on the number of clusters, which is a required input parameter, is recommended to be chosen slightly superior to the number of classes.

When the algorithm converges, the partitioning of the set of image pixels into clusters is obtained. However, as no spatial information is used during the clustering procedure, pixels with the same cluster label can form a connected spatial region, or can belong to disjoint regions. In order to obtain a segmentation map, a connected components labeling algorithm [28] is applied to the output image partitioning obtained by clustering.

The total number of parameters to be estimated by the EM algorithm is $P = (B(B + 1)/2 + B + 1)C + 1$, where B is a dimensionality of feature vectors. If the value of B is large, P may be quite a large number. This may cause the problem of the covariance matrix singularity or inaccurate parameter estimation results. In order to avoid these problems, we propose to previously apply a feature reduction, using the method of piecewise constant function approximations (PCFA) [29], which has shown good performances for hyperspectral data feature extraction.

C. RHSEG segmentation

The Hierarchical image SEGmentation (HSEG) algorithm is a segmentation technique based on iterative hierarchical step-wise optimization region growing method. Furthermore, it provides a possibility of merging non-adjacent regions by spectral clustering [30].

The following outline of the HSEG algorithm is based on the description given in [31], [30]:

- 1) Initialize the segmentation by assigning for each pixel a region label. If a pre-segmentation is provided, label each pixel according to the pre-segmentation. Otherwise, label each pixel as a separate region.
- 2) Calculate the dissimilarity criterion value between all pairs of spatially adjacent regions.
- 3) Find the smallest dissimilarity criterion value *dissim_val* and set *thresh_val* equal to it. Then

merge all pairs of spatially adjacent regions with $dissim_val = thresh_val$.

- 4) If a parameter $S_{wght} > 0.0$, merge all pairs of spatially non-adjacent regions with $dissim_val \leq S_{wght} \cdot thresh_val$.
- 5) If convergence is not achieved, go to step (2).

In order to reduce computational demands, a Recursive divide-and-conquer approximation of HSEG (RHSEG) has been developed. The NASA-Goddard RHSEG software provides an efficient implementation of the RHSEG algorithm.

When determining most similar pair of regions, we propose to choose the standard Spectral Angle Mapper (SAM) between the region mean vectors and as the dissimilarity criterion [30]. The SAM measure between \mathbf{u}_i and \mathbf{u}_j ($\mathbf{u}_i, \mathbf{u}_j \in \mathbb{R}^B$) determines the spectral similarity between two vectors by computing the angle between them. It is defined as

$$SAM(\mathbf{u}_i, \mathbf{u}_j) = \arccos \left(\frac{\sum_{b=1}^B u_{ib}u_{jb}}{[\sum_{b=1}^B u_{ib}^2]^{1/2}[\sum_{b=1}^B u_{jb}^2]^{1/2}} \right). \quad (5)$$

The optional parameter S_{wght} tunes the relative importance of spectral clustering *versus* region growing. If $S_{wght} = 0.0$, only merging of spatially adjacent regions is performed. If $0.0 < S_{wght} \leq 1.0$, merging between spatially adjacent regions is favored compared to merging of spatially non-adjacent regions by a factor of $1.0/S_{wght}$. As discussed in [32], the optimal parameter S_{wght} can be chosen based on *a priori* knowledge about information classes contained in the image. If some classes have very similar spectral responses, we recommended to choose $S_{wght} = 0.0$, i.e., to perform segmentation only in the spatial domain. Otherwise, we recommend to include the possibility of merging spatially non-adjacent regions, while favoring region growing (for instance, $S_{wght} = 0.1$ can be chosen). If $S_{wght} > 0.0$, labeling of connected components has to be applied after RHSEG in order to obtain a segmentation map where each spatially connected component has a unique label.

RHSEG provides as output a hierarchical sequence of image partitions. In this sequence, a particular object can be represented by several regions at finer levels of details, and can be assimilated with other objects in one region at coarser levels of details. While this hierarchical sequence allows flexibility in choosing the appropriate level of detail for the segmentation map, the manual selection of the appropriate level can be time consuming. Therefore, automatic selection is desirable. The possibility of the automated selection of results for the RHSEG was explored in [33].

D. Pixelwise classification

Independent of the previous steps, a pixelwise classification of the hyperspectral image is performed. We propose to use an SVM classifier for this purpose. Other pixelwise classifier could be used. However, SVMs perform extremely well in classifying high-dimensional data when a limited number of training samples are available [5], [34]. We refer the reader to [5], [35] for details on SVM technique. This step results in a classification map, where each pixel has a unique class label.

E. Majority voting within segmentation regions

Each of the obtained unsupervised segmentation maps is combined with the pixelwise classification map using the majority voting principle: For every region in the segmentation map, all the pixels are assigned to the most frequent class within this region (see an illustrative example in Figure 2). Thus, q segmentation maps combined with the pixelwise classification map result in q spectral-spatial classification maps (since we propose to use three different segmentation techniques, in this particular case $q = 3$).

F. Marker selection

This step consists of computing a map of markers, using spectral-spatial classification maps from the previous step and exclusionary rule: For every pixel, if all the classifiers agree, the pixel is kept as a marker, with the corresponding class label. The resulting map of m markers contains the most reliably classified pixels.

G. Construction of a Minimum Spanning Forest

In the final step, image pixels are grouped into an MSF rooted on the selected markers [15]. Each pixel is considered as a vertex $v \in V$ of an undirected graph $G = (V, E, W)$, where V and E are the sets of vertices and edges, respectively, and W is a mapping of the set of the edges E into \mathbb{R}^+ . Each edge $e_{i,j} \in E$ of this graph connects a couple of vertices i and j corresponding to the neighboring pixels. Furthermore, a weight $w_{i,j}$ is assigned to each edge $e_{i,j}$, which indicates the degree of dissimilarity between two vertices (i.e., two corresponding pixels) connected by this edge. We propose to use an 8-neighborhood and the SAM measure for computing weights of edges, as described in [15].

Given a graph $G = (V, E, W)$, the MSF rooted on a set of m distinct vertices $\{t_1, \dots, t_m\}$ consists in finding a spanning forest $F^* = (V, E_{F^*})$ of G , such that each distinct tree of F^* is grown from one root t_i , and the sum of the edges weights of F^* is minimal [36]

$$F^* \in \arg \min_{F \in SF} \left\{ \sum_{e_{i,j} \in E_F} w_{i,j} \right\}, \quad (6)$$

where SF is a set of all spanning forests of G rooted on $\{t_1, \dots, t_m\}$.

Algorithm 1 Prim's Algorithm

Require: Connected graph $G = (V, E, W)$

Ensure: Tree $T^* = (V^*, E^*, W^*)$

$V^* = \{v\}$, v is an arbitrary vertex from V

while $V^* \neq V$ **do**

Choose edge $e_{i,j} \in E$ with minimal weight such that

$i \in V^*$ and $j \notin V^*$

$V^* = V^* \cup \{j\}$

$E^* = E^* \cup \{e_{i,j}\}$

end while

In order to obtain the MSF rooted on m markers corresponding to the vertices $t_i, i = 1, \dots, m$, an additional

root vertex r is added and is connected by the null-weight edges to the vertices t_i . The minimum spanning tree of the constructed graph induces an MSF in G , where each tree is grown on a vertex t_i ; the MSF is obtained after removing the vertex r . Prim's algorithm can be used for building the MSF (see Algorithm 1) [37]. The efficient implementation of the algorithm using a binary min-heap is possible [38], the resulting time complexity of the algorithm is $O(|E|\log|V|)$. Finally, a spectral-spatial classification map is obtained by assigning the class of each marker to all the pixels grown from this marker.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Two different hyperspectral images were used for the experiments, with different contexts (one urban area and one agricultural area) and recorded by different sensors (RODIS and AVIRIS airborne imaging spectrometers). These data sets and the corresponding results are discussed in the next two sections.

A. Classification of the University of Pavia Image

The *University of Pavia* image was recorded by the RODIS optical sensor over the urban area of the University of Pavia, Italy. The image is 610×340 pixels, with a spatial resolution of 1.3 m/pixel. The number of data channels in the acquired image is 115 (with a spectral range from 0.43 to 0.86 μm). The 12 most noisy channels have been removed, and the remaining 103 bands were used for the experiments. Nine classes of interest are considered, which are detailed in Table I. Figure 5 shows a three-band false color image and the reference data. More information about the image, with the number of training and test samples for each class can be found in [12].

The segmentation of the considered image was performed, using the three different techniques discussed in the previous section. For the EM algorithm, the maximum number of clusters was chosen equal to 10 (typically slightly superior to the number of classes). Before applying the EM technique, a feature extraction on the original 103-band image was applied, using the method of PCFA [29] to get a 10-band image \mathbf{Y}_{UP} . Pixels from the training set have been used for selecting features. The method has averaged the following groups of adjacent spectral channels: 1-4, 5-10, 11-24, 25-35, 36-43, 44-68, 69-72, 73-75, 76-79, 80-103.

For the RHSEG algorithm, we have chosen $S_{wght} = 0.1$, since the image of an urban area contains classes with mostly dissimilar spectral responses. Furthermore, a segmentation map at the relevant level of hierarchy was chosen interactively with the program HSEGVier [30]. The obtained watershed, EM and RHSEG segmentation maps contained 11802, 22549 and 7575 regions, respectively.

The multiclass pairwise SVM classification, with the Gaussian Radial Basis Function (RBF) kernel, of the original image was performed, with the parameters chosen by fivefold cross validation: $C = 128$, $\gamma = 0.125$. The results of the pixelwise classification were combined with the segmentation results, using the majority voting approach. Finally, the marker selection (see Figure 5(g); 132521, i.e., 64% of pixels were selected

as markers) and the construction of an MSF were performed, resulting in the *MSSC-MSF* spectral-spatial classification map depicted in Figure 5(h).

Table I summarizes the global and class-specific accuracies of the pixelwise SVM, segmentation plus majority voting (*WH+MV*, *EM+MV*, *RHSEG+MV* for three segmentation techniques, respectively) and the proposed *MSSC-MSF* classification methods. The following measures of accuracy were used: Overall Accuracy (OA is the percentage of correctly classified pixels), Average Accuracy (AA is the mean of class-specific accuracies, i.e., the percentage of correctly classified pixels for each class) and kappa coefficient (κ is the percentage of agreement, i.e., correctly classified pixels, corrected by the number of agreements that would be expected purely by chance [39]). In order to compare performances of the proposed technique with the previously proposed methods, we have also included results of the well-known ECHO spatial classifier [10], as well as the results obtained using the construction of an MSF from the probabilistic SVM-derived markers followed by majority voting within connected regions (*SVMMSF+MV*) [15].

Furthermore, we assess the importance of spectral-spatial approaches for marker selection. For this purpose, we have replaced the *WH+MV*, *EM+MV*, *RHSEG+MV* classification maps by three maps obtained using standard pixelwise classification techniques (we call this modified scheme an *MC-MSF* classification method). SVM, Maximum Likelihood (*ML*) and 3-Nearest Neighborhood (*3-NN*, using the SAM distance) methods have been used for this purpose. The *ML* and the *3-NN* techniques were applied on the 10-band image \mathbf{Y}_{UP} feature vectors. The accuracies of the modified *MC-MSF* classification, as well as pixelwise classification results are given in Table I.

As can be seen from Table I, the SVM method gives the highest accuracies among all the pixelwise classification techniques. All the spectral-spatial approaches yield higher classification accuracies when compared to pixelwise methods. The proposed MC approach for marker selection improves accuracies, when compared to those obtained by classification techniques used in the preliminary step of the marker selection procedure, both for the *MC-MSF* and *MSSC-MSF* methods. The best global and the best class-specific accuracies for most classes are achieved by applying the proposed *MSSC-MSF* method. In this case, the overall accuracy is improved by 16.9 percentage points and the average accuracy is improved by 10.3 percentage points, when compared to the SVM classification. All the class-specific accuracies are higher than 96%. Only the accuracy for the class *shadows*, representing small spatial structures, is slightly reduced when compared to the SVM results (the drawback of applying spectral-spatial classification to small structures was discussed for instance in [12]). The *MSSC-MSF* classification accuracies are much higher than the *MC-MSF* accuracies. Furthermore, the presented classification accuracies are higher than all previous results we have found in the literature for the same data.

Figure 5 depicts the *MC-MSF* and *MSSC-MSF* classification maps, as well as the SVM, *ECHO* and *SVMMSF+MV* classification maps given for comparison. In Figure 5(g) [*MSSC-*

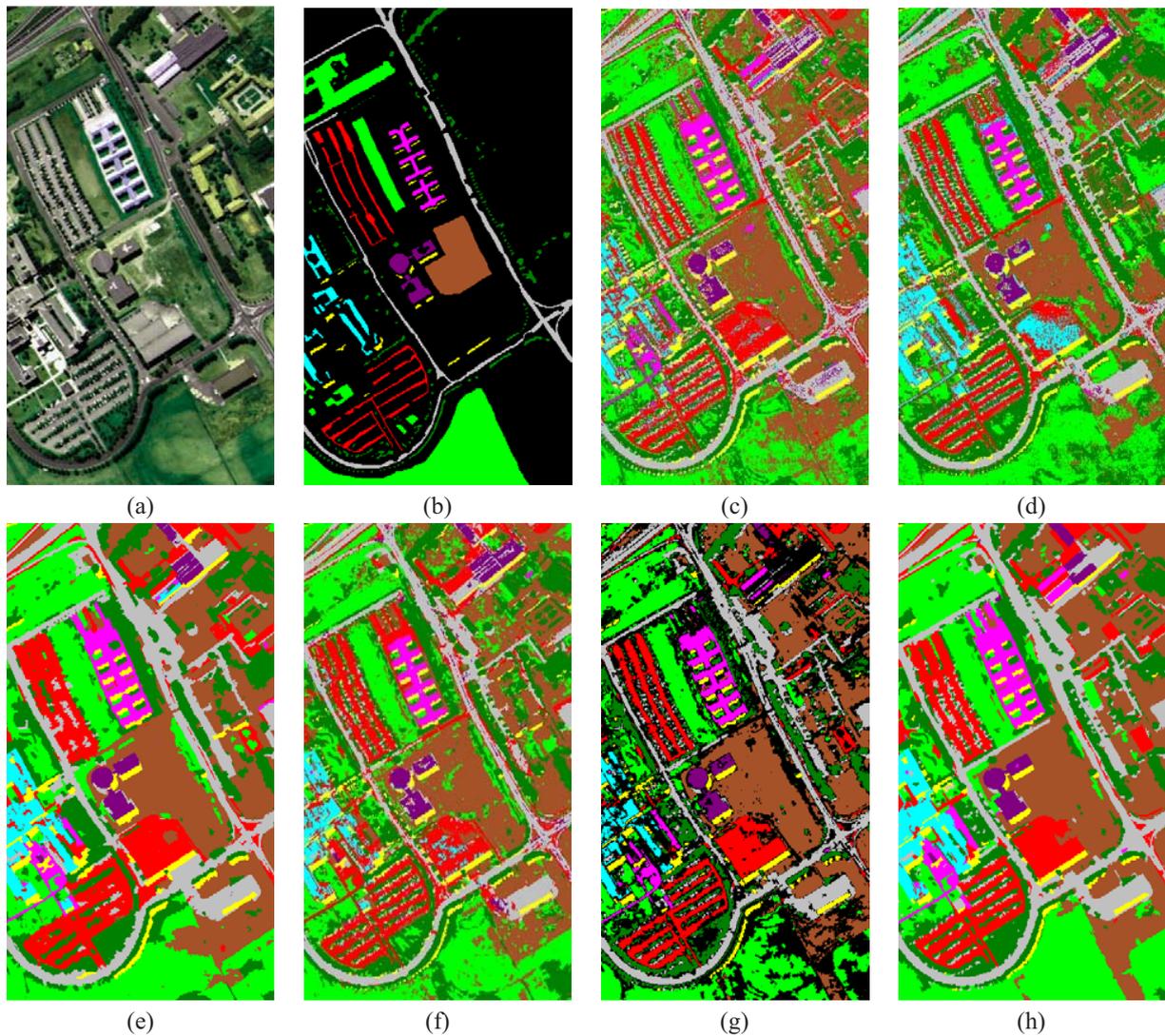


Fig. 5. *University of Pavia* image. (a) Three-band color composite. (b) Reference data: Asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks and shadows. (c) SVM pixelwise classification map. (d) *ECHO* classification map. (e) *SVMMSF+MV* classification map. (f) *MC-MSF* classification map. (g) *MSSC-MSF* marker map. (h) *MSSC-MSF* classification map.

TABLE I
CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE *University of Pavia* IMAGE: OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), KAPPA COEFFICIENT (κ) AND CLASS-SPECIFIC ACCURACIES.

	3-NN	ML	SVM	ECHO	WH+MV	EM+MV	RHSEG +MV	SVMSF +MV	MC- MSF	MSSC- MSF
OA	68.38	79.06	81.01	87.58	85.42	94.00	93.85	91.08	87.98	97.90
AA	77.21	84.85	88.25	92.16	91.31	93.13	97.07	94.76	92.05	98.59
κ	59.85	72.90	75.86	83.90	81.30	91.93	91.89	88.30	84.32	97.18
Asphalt	64.96	76.43	84.93	87.98	93.64	90.10	94.77	93.16	87.01	98.00
Meadows	63.18	75.99	70.79	81.64	75.09	95.99	89.32	85.65	83.24	96.67
Gravel	62.31	64.57	67.16	76.91	66.12	82.26	96.14	89.15	75.37	97.80
Trees	95.95	97.08	97.77	99.31	98.56	85.54	98.08	91.24	98.97	98.83
Metal sheets	99.73	99.91	99.46	99.91	99.91	100	99.82	99.91	99.91	99.91
Bare soil	57.42	70.03	92.83	93.96	97.35	96.72	99.76	99.91	93.24	100
Bitumen	82.67	90.62	90.42	92.97	96.23	91.85	100	98.57	95.11	99.90
Bricks	77.08	90.10	92.78	97.35	97.92	98.34	99.29	99.05	97.00	99.76
Shadows	91.57	98.87	98.11	99.37	96.98	97.36	96.48	96.23	98.62	96.48

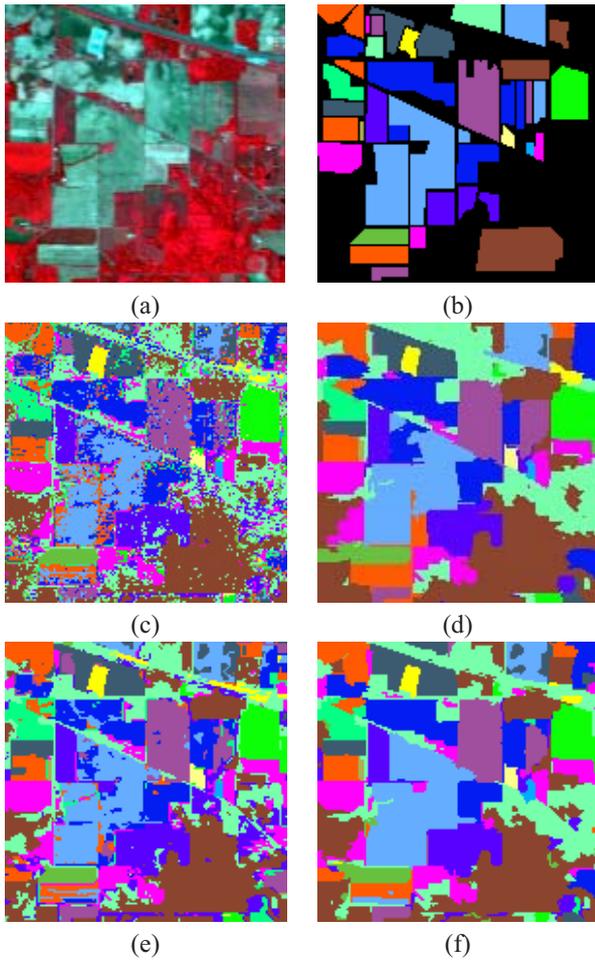


Fig. 6. *Indian Pines* image. (a) Three-band color composite. (b) Reference data: Corn-no till, Corn-min till, Corn, Soybeans-no till, Soybeans-min till, Soybeans-clean till, Alfalfa, Grass/pasture, Grass/trees, Grass/pasture-mowed, Hay-windrowed, Oats, Wheat, Woods, Bldg-Grass-Tree-Drives, Stone-steel towers. (c) SVM pixelwise classification map. (d) *SVMSF+MV* classification map. (e) *MC-MSF* classification map. (f) *MSSC-MSF* classification map.

MSF marker map], it can be seen that the marker pixels, i.e., the most reliable classified pixels, are typically located at the center of spatial structures, while borders of structures are under a high risk of being misclassified. The *MSSC-MSF* classification map [see Figure 5(h)] contains much more homogeneous regions when compared to the maps obtained by other pixelwise and spectral-spatial approaches. These results prove the importance of the use of MC systems and spatial information throughout the classification procedure.

B. Classification of the *Indian Pines* Image

The proposed scheme was also tested on the *Indian Pines* image of a vegetation area, acquired by the AVIRIS sensor over the Indian Pines site in Northwestern Indiana. The image has spatial dimensions of 145 by 145 pixels, with a spatial resolution of 20 m/pixel. Twenty water absorption bands (104-108, 150-163, 220) have been removed [16], and a 200-band image was used for the experiments. The reference data contain sixteen classes of interest, which represent mostly

different types of crops and are detailed in Table II. A three-band false color image and the reference data are presented in Figure 6. We have randomly chosen 50 samples for each class from the reference data as training samples, except for classes “*alfalfa*”, “*grass/pasture-mowed*” and “*oats*”. These classes contain a small number of samples in the reference data. Therefore, only 15 samples for each of these classes were chosen randomly to be used as training samples. The remaining samples composed the test set.

Segmentation of the *Indian Pines* image was performed, using the three discussed techniques. For the EM technique, the upper bound on the number of classes was chosen equal to 17, and a feature reduction has been previously applied, using the method of PCFA [29] to get a 10-band image Y_{IN} . The following groups of band have been averaged: 1-18, 19-36, 37-53, 54-57, 58-61, 62-75, 76-81, 82-99, 100-140, 141-200.

Since some classes have very similar spectral responses in the *Indian Pines* image (for instance, three classes of corn and three classes of soybeans), we have set $S_{weight} = 0.0$ for the RHSEG method. A segmentation map at the relevant level of hierarchy was chosen with the program HSEGVIEWER. The resulting watershed, EM and RHSEG segmentation maps contained 1277, 3832 and 823 regions, respectively.

A pixelwise classification on the 200-band image was performed, using the multiclass one *versus* one SVM classifier with the Gaussian RBF kernel. The optimal parameters C and γ were chosen by fivefold cross validation: $C = 128$, $\gamma = 2^{-6}$. After the segmentation results have been combined with the pixelwise classification map, the marker selection (14409, i.e., 68% of pixels were selected as markers) and the MSF construction were applied, as described in the previous section.

Table II gives the global and class-specific accuracies of the pixelwise SVM, segmentation followed by majority voting and the proposed *MSSC-MSF* classification techniques. The performances of the proposed approach are compared with those obtained by the ECHO and *SVMSF+MV* methods, as described in the previous section. Finally, the *MC-MSF* classification was applied in the same way as for the previous data set.

From the table, the same conclusions as for the previous data set can be derived. The SVM classification yields the best accuracies among all the applied pixelwise methods. Spectral-spatial classification accuracies are always higher when compared to pixelwise accuracies. The proposed MC method succeeds in combining several classification results for further improvement of accuracies. The *MSSC-MSF* yields the best OA, kappa coefficient and most of the class-specific accuracies. The AA is only slightly (non-significantly) lower when compared to that obtained by the recently proposed *SVMSF+MV* method. The *MSSC-MSF* results outperform the *MC-MSF* ones.

Figure 6 shows the SVM, *MC-MSF*, *MSSC-MSF* and *SVMSF+MV* classification maps. As can be seen, the *MSSC-MSF* map contains much more homogeneous spatial structures, when compared to the SVM and *MC-MSF* maps, and is comparable with the *SVMSF+MV* map. Again, spectral-spatial marker-based techniques yielded the most accurate

TABLE II
CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE *Indian Pines* IMAGE: OVERALL ACCURACY (OA), AVERAGE ACCURACY (AA), KAPPA COEFFICIENT (κ) AND CLASS-SPECIFIC ACCURACIES.

	3-NN	ML	SVM	ECHO	WH+MV	EM+MV	RHSEG +MV	SVMMMSF +MV	MC- MSF	MSSC- MSF
OA	66.27	75.41	78.17	82.64	86.63	83.60	90.86	91.80	86.66	92.32
AA	76.77	79.61	85.97	83.75	91.61	85.34	93.96	94.28	92.58	94.22
κ	62.04	72.25	75.33	80.38	84.83	81.43	89.56	90.64	84.82	91.19
Corn-no till	41.84	71.39	78.18	83.45	94.22	89.09	90.46	93.21	83.82	89.74
Corn-min till	62.24	63.01	69.64	75.13	78.06	75.64	83.04	96.56	74.62	86.99
Corn	73.37	85.87	91.85	92.39	88.59	65.22	95.65	95.65	96.74	95.11
Soybeans-no till	67.43	79.43	82.03	90.10	96.30	88.14	92.06	93.91	93.36	91.84
Soybeans-min till	53.91	52.65	58.95	64.14	68.82	65.67	84.04	81.97	72.91	89.16
Soybeans-clean till	64.72	85.99	87.94	89.89	90.78	95.04	95.39	97.16	95.92	97.34
Alfalfa	84.62	48.72	74.36	48.72	94.87	94.87	92.31	94.87	94.87	94.87
Grass/pasture	86.35	93.51	92.17	94.18	95.08	93.96	94.41	94.63	98.21	94.63
Grass/trees	91.97	94.69	91.68	96.27	97.99	96.41	97.56	97.27	97.70	97.85
Grass/pasture-mowed	100	36.36	100	36.36	100	100	100	100	100	100
Hay-windrowed	95.67	97.72	97.72	97.72	99.54	99.32	99.54	99.77	99.54	99.77
Oats	80.00	100	100	100	100	40.00	100	100	100	100
Wheat	99.38	98.15	98.77	98.15	99.38	98.77	98.15	99.38	99.38	99.38
Woods	86.17	95.42	93.01	94.21	97.11	96.70	98.63	99.68	98.47	99.44
Bldg-Grass-Tree-Drives	45.15	73.03	61.52	81.52	69.39	66.67	82.12	68.79	77.88	73.64
Stone-steel towers	95.56	97.78	97.78	97.78	95.56	100	100	95.56	97.78	97.78

classification maps.

Although for the *Indian Pines* image, the *MSSC-MSF* and *SVMMMSF+MV* methods yield similar results, here we stress the advantages of the proposed *MSSC-MSF* approach *versus* the previous one for spectral-spatial classification:

- **Robustness:** While for the *SVMMMSF+MV* method the marker selection strongly depends on the performances of the selected pixelwise classifier, the MC approach mitigates this dependence. Since in the *MSSC-MSF* scheme, different segmentation maps are combined with one pixelwise classification map, the choice of the classifier is also important. However, if in the *SVMMMSF+MV* method a pixel was wrongly classified with a high probability, it will yield a wrong marker. In the new approach, the majority voting within segmentation regions can correct the misclassification result for a particular pixel, before the marker map is built.
- **Computational Complexity:** In the *SVMMMSF+MV* method, the probabilistic pixelwise SVM classification part is the most time-consuming [40]. In the *MSSC-MSF* approach, SVM classification is performed without the computation of probability estimates; this reduces the pixelwise classification part execution time. The unsupervised segmentation techniques are much less time-consuming, when compared to the SVM-classification. Furthermore, their efficient implementations are available, and they can be executed in parallel at the same time with the SVM classification. As a conclusion, the efficient implementation of the proposed *MSSC-MSF* approach is possible, which would run faster than the previously proposed *MSSC-MSF* method.

V. CONCLUSIONS

Hyperspectral sensors capture images in hundreds of narrow spectral channels. The detailed spectral signatures for each

spatial location provide rich information about an image scene, leading to better discrimination between physical materials and objects. However, interpretation of these high-dimensional signatures is a challenging task. Although pixelwise classification techniques have given high classification accuracies when dealing with hyperspectral data, the incorporation of the spatial context into classification procedures yields further improvement of the accuracies.

In this paper, a new method for spectral-spatial classification of hyperspectral images based on multiple classifiers is proposed. First, a marker map is constructed by selecting the pixels assigned by several spectral-spatial classifiers to the same class. This ensures a robust and reliable selection. Then, an MSF rooted on the selected markers is built. Experimental results demonstrated that the proposed method improves classification accuracies, when compared to previously proposed classification schemes, and provides classification maps with homogeneous regions.

In conclusion, the proposed methodology succeeded in taking advantage of multiple classifiers and the spatial and the spectral information simultaneously for accurate hyperspectral image classification. The method performs well for different data sets, i.e., data containing large spatial structures and/or small and complex structures, with spectrally dissimilar and/or spectrally confusing classes. Finally, the efficient implementation of the proposed approach is possible. In the future, we will further investigate the integration of spectral-spatial approaches in MC systems for accurate and robust classification of hyperspectral images.

ACKNOWLEDGMENTS

The authors would like to thank P. Gamba from the University of Pavia, Italy, and D. Landgrebe from Purdue University, USA, for providing the hyperspectral data.

REFERENCES

- [1] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*. Wiley-Interscience, 2007.
- [2] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. S. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the Airborne Visible Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, September 1998.
- [3] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley & Sons, Inc., 2003.
- [4] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn," *Comput. Electron. Agricult.*, vol. 39, pp. 67–93, 2003.
- [5] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geos. and Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, June 2005.
- [6] M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data," Ph.D. dissertation, Grenoble Institute of Technology, 2007.
- [7] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geos. and Remote Sens.*, vol. 46, no. 10, Oct. 2008.
- [8] A. Farag, R. Mohamed, and A. El-Baz, "A unified framework for map estimation in remote sensing image segmentation," *IEEE Trans. Geos. and Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, July 2005.
- [9] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Classification of hyperspectral data using Support Vector Machines and adaptive neighborhoods," in *Proc. of the 6th EARSeL SIG IS workshop*, Tel Aviv, Israel, 2009.
- [10] R. L. Kettig and D. A. Landgrebe, "Classification of multispectral image data by extraction and classification of homogeneous objects," *IEEE Trans. Geoscience Electronics*, vol. 14, no. 1, pp. 19–26, Jan. 1976.
- [11] R. Gonzalez and R. Woods, *Digital Image Processing, Second Edition*. Prentice Hall, 2002.
- [12] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitioned clustering techniques," *IEEE Trans. Geos. and Remote Sens.*, vol. 47, no. 9, pp. 2973–2987, Sept. 2009.
- [13] Y. Tarabalka, J. Chanussot, J. A. Benediktsson, J. Angulo, and M. Fauvel, "Segmentation and classification of hyperspectral data using watershed," in *Proc. of IGARSS '08*, Boston, USA, 2008, pp. III-652 – III-655.
- [14] P. Soille, *Morphological Image Analysis*, 2nd ed. Springer-Verlag, 2003.
- [15] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers," *IEEE Trans. Systems, Man, and Cybernetics: Part B*, in press.
- [16] S. Tadjudin and D. A. Landgrebe, "Covariance estimation with limited training samples," *IEEE Trans. Geos. and Remote Sens.*, vol. 37, no. 4, pp. 2113–2118, July 1999.
- [17] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Systems, Man and Cybernetics*, vol. 22, no. 3, pp. 418–435, May/June 1992.
- [18] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, Mar. 1998.
- [19] G. Briem, J. A. Benediktsson, and J. R. Sveinsson, "Multiple classifiers applied to multisource remote sensing data," *IEEE Trans. Geos. and Remote Sens.*, vol. 40, no. 10, pp. 2291–2299, Oct. 2002.
- [20] G. Fumera and F. Roli, "A theoretical and experimental analysis of linear combiners for multiple classifier systems," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 942–956, June 2005.
- [21] Y.-C. Tzeng, "Remote sensing images classification/data fusion using distance weighted multiple classifier systems," in *Proc. of PDCAT'06*, Taipei, Taiwan, Dec. 2006.
- [22] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," in *Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*. Rennes, France: CCETT/IRISA, Sept. 1979.
- [23] G. Noyer, J. Angulo, and D. Jeulin, "Morphological segmentation of hyperspectral images," *Image Analysis & Stereology*, vol. 26, pp. 101–109, 2007.
- [24] A. Evans and X. Liu, "A morphological gradient approach to color edge detection," *IEEE Trans. Image Processing*, vol. 15, no. 6, pp. 1454–1463, June 2006.
- [25] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Analysis and Machine Intel.*, vol. 13, no. 6, pp. 583–598, June 1991.
- [26] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *Proc. of the IEEE*, vol. 78, no. 4, pp. 678–689, 1990.
- [27] P. Masson and W. Pieczynski, "SEM algorithm and unsupervised segmentation of satellite images," *IEEE Trans. Geos. and Remote Sens.*, vol. 31, no. 3, pp. 618–633, 1993.
- [28] L. Shapiro and G. Stockman, *Computer Vision*. Prentice Hall, 2002.
- [29] A. Jensen and A. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geos. and Remote Sens. Letters*, vol. 4, no. 4, pp. 547–551, Oct. 2007.
- [30] J. C. Tilton, "HSEG/RHSEG, HSEGViewer and HSEGReader user's manual (version 1.40)," *Provided with the evaluation version of RHSEG available from: <http://ipp.gsfc.nasa.gov/RHSEG>*, 2008.
- [31] J. Tilton, "Analysis of hierarchically related image segmentations," *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 60–69, 2003.
- [32] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral images using segmentation-derived adaptive neighborhoods," in *Multivariate Image Processing, ISTE Ltd and John Wiley & Sons Inc*, 2009, pp. 341–374.
- [33] A. J. Plaza and J. C. Tilton, "Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images," in *Proc. of IGARSS '05*, vol. 7, July 2005, pp. 4946–4949.
- [34] G. Licciardi, F. Pacifici, D. Tuia, S. Prasad, T. West, F. Giacco, J. Inglada, E. Christophe, J. Chanussot, and P. Gamba, "Decision fusion for the classification of hyperspectral data: Outcome of the 2008 GRS-S data fusion contest," *IEEE Trans. Geos. and Remote Sens.*, vol. 47, no. 11, pp. 3857–3865, Nov. 2009.
- [35] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
- [36] J. Stawiński, "Mathematical morphology and graphs: Application to interactive medical image segmentation," Ph.D. dissertation, Paris School of Mines, 2008.
- [37] R. C. Prim, "Shortest connection networks and some generalizations," *Bell System Technology Journal*, vol. 36, pp. 1389–1401, 1957.
- [38] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, Second Edition*. MIT Press and McGraw-Hill, 2001.
- [39] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag New York, Inc., 1999.
- [40] N. Cristianini and J. Shawe-Taylor, *Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.



Yuliya Tarabalka (S'08) received the B.S. degree in computer science from Ternopil Ivan Pul'uj State Technical University, Ukraine, in 2005 and M.Sc. degree in signal and image processing from the Grenoble Institute of Technology (INPG), France, in 2007. She is currently working toward the Ph.D. degree at the University of Iceland, Reykjavik, Iceland, and at INPG.

From July 2007 to January 2008, she was a Researcher with the Norwegian Defence Research Establishment, Norway. Her current research work is funded by the "HYPER-I-NET" Marie Curie Research Training Network. Her research interests are in the areas of image processing, pattern recognition, hyperspectral imaging and development of efficient algorithms.



Jón Atli Benediktsson (S'84M'90SM'99F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, in 1984, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively. He is currently Pro Rector for Academic Affairs and Professor of Electrical and Computer Engineering at the University of Iceland. His research interests are in remote sensing, pattern recognition, neural networks, image processing, and signal processing, and he has published extensively in those fields. Dr. Benediktsson was Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGARS) from 2003 to 2008. He is a Fellow of the IEEE. He has been on the AdCom of the IEEE Geoscience and Remote Sensing Society from 2000 and is the GRSS Executive Vice President for 2010-2011. He received the Stevan J. Kristof Award from Purdue University in 1991 as outstanding graduate student in remote sensing. In 1997, Dr. Benediktsson was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award, in 2000, he was granted the IEEE Third Millennium Medal, in 2004, he was a co-recipient of the University of Iceland's Technology Innovation Award, in 2006 he received the yearly research award from the Engineering Research Institute of the University of Iceland, and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a member of Societas Scinetiarum Islandica and Tau Beta Pi.



James C. Tilton (S'79-M'81-SM'94) received B.A. degrees in electronic engineering, environmental science and engineering, and anthropology and a M. E. E. (electrical engineering) from Rice University, Houston, TX in 1976. He also received an M. S. in optical sciences from the University of Arizona, Tucson, AZ in 1978 and a Ph. D. in electrical engineering from Purdue University, West Lafayette, IN in 1981.

He is currently a Computer Engineer with the Computational and Information Sciences and Technology Office (CISTO) of the Science and Exploration Directorate at the NASA Goddard Space Flight Center, Greenbelt, MD. As a member of the CISTO, Dr. Tilton is responsible for designing and developing computer software tools for space and earth science image analysis, and encouraging the use of these computer tools through interactions with space and earth scientists. His development of a recursive hierarchical segmentation algorithm has resulted in a patent and two other patent applications.

Dr. Tilton is a senior member of the IEEE Geoscience and Remote Sensing and Signal Processing Societies, and is a member of Phi Beta Kappa, Tau Beta Pi and Sigma Xi. From 1992 through 1996, he served as a member of the IEEE Geoscience and Remote Sensing Society Administrative Committee. Since 1996 he has served as an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing.



Jocelyn Chanussot (M'04SM'04) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA-French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab, INPG). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. Dr. Chanussot is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is the Cochair of the GRS Data Fusion Technical Committee (2005-2008) and a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008). He is the founding President of IEEE GRS French chapter (2007).

Paper 7

Y. Tarabalka, T. V. Haavardsholm, I. Kasen, and T. Skauli, “Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing,” *Journal of Real-Time Image Processing*, vol. 4, no. 3, pp. 287-300, Aug. 2009.

Real-time anomaly detection in hyperspectral images using multivariate normal mixture models and GPU processing

Yuliya Tarabalka · Trym Vegard Haavardsholm ·
Ingebjørg Kåsen · Torbjørn Skauli

Abstract Hyperspectral imaging, which records a detailed spectrum of light arriving in each pixel, has many potential uses in remote sensing as well as other application areas. Practical applications will typically require real-time processing of large data volumes recorded by a hyperspectral imager. This paper investigates the use of graphics processing units (GPU) for such real-time processing. In particular, the paper studies a hyperspectral anomaly detection algorithm based on normal mixture modelling of the background spectral distribution, a computationally demanding task relevant to military target detection and numerous other applications. The algorithm parts are analysed with respect to complexity and potential for parallelization. The computationally dominating parts are implemented on an Nvidia GeForce 8800 GPU using the Compute Unified Device Architecture programming interface. GPU computing performance is compared to a multi-core central processing unit implementation. Overall, the GPU implementation runs significantly faster, particularly for highly data-parallelizable and arithmetically intensive algorithm parts. For the parts related to covariance computation, the speed gain is less pronounced, probably due to a smaller ratio of arithmetic to memory access. Detection results on an actual data set demonstrate that the total

speedup provided by the GPU is sufficient to enable real-time anomaly detection with normal mixture models even for an airborne hyperspectral imager with high spatial and spectral resolution.

Keywords Anomaly detection · Hyperspectral imagery · Multivariate normal mixture model · General purpose GPU processing

1 Introduction

Hyperspectral imaging is characterized by its ability to record detailed information about the spectral distribution of the received light. Hyperspectral imaging sensors typically measure the energy of the received light in tens or hundreds of narrow spectral bands in each spatial position in the image, so that each pixel in a hyperspectral image can be represented as a high-dimensional vector containing the sampled spectrum. Since different substances exhibit different spectral signatures, hyperspectral imaging is a well-suited technology for numerous remote sensing applications including target detection.

When no information about the spectral signature of the desired targets is available, a popular approach for target detection is to look for objects that deviate from the typical spectral characteristics in the image. This approach is commonly referred to as anomaly detection [17], and is related to what is often called outlier detection in statistics. If targets are small compared to the image size, the spectral characteristics in the image are dominated by the background. An important step in anomaly detection is therefore often to compute a metric for correspondence with the background, which then can be thresholded to detect objects that are unlikely to be background objects.

Y. Tarabalka · T. V. Haavardsholm · I. Kåsen · T. Skauli (✉)
Norwegian Defence Research Establishment (FFI),
P.O. Box 25, 2007 Kjeller, Norway
e-mail: torbjorn.skauli@ffi.no

Y. Tarabalka
e-mail: Yuliya.Tarabalka@gipsa-lab.inpg.fr

T. V. Haavardsholm
e-mail: trym.haavardsholm@ffi.no

I. Kåsen
e-mail: ingebjorg.kasen@ffi.no

Hyperspectral imaging inherently produces large volumes of data which create challenges in data transfer, storage and processing. In particular, real-time processing of hyperspectral imagery is no trivial task. Nevertheless, it is highly desirable in target detection and other applications to process images in real time, usually on board the platform carrying the sensor.

Several real-time anomaly detection methods suitable for on-board processing exist, like the SSRX implemented in the ARCHER and WAR HORSE programs [18, 19], but these are usually based on very simple geometric or statistical representations of the image background variability. In contrast, mixture models, such as the multivariate normal mixture model, may be able to represent the background variability quite accurately, resulting in statistically meaningful background metrics. The characteristics of anomaly detection based on normal mixture models are discussed in some detail in [5]. This anomaly detector has demonstrated good detection performance on several occasions. One of the main criticisms of this method, however, has been that it is computationally very expensive, and therefore poorly suited for on-board real-time target detection.

Fortunately, some of the most time-consuming tasks in the normal mixture model processing are easily parallelized, so that the multi-core architecture in modern central processing units (CPUs) may be exploited to speed up the processing. An interesting recent development has been the introduction of fully programmable graphics processing units (GPUs) together with software interfaces like NVIDIA CUDA [12] and AMD CTM [1] dedicated to general purpose processing on video cards. Because the GPU architectures are optimized for massively parallel processing, modern commodity video cards can achieve very high computational performance for parallel problems, peaking at several hundred GFLOPS or more. The high demand for realistic graphics (and physics) in the computer game market drives the development of increasingly powerful GPUs at low cost, while keeping computer architectures adapted to this technology to achieve very high bandwidth communication between the computer and the graphics hardware. Today, low-cost, low-weight gaming computers are readily available with extremely powerful parallel computing performance. This kind of hardware is therefore very well suited for on-board processing in a hyperspectral target detection scenario.

Although general-purpose computing on graphics processing units (GPGPU) has been an active area of research for decades, the introduction of Compute Unified Device Architecture (CUDA) and CTM has finally brought it within reach of a broader community, giving programmers access to dedicated application programming interfaces (APIs), software development kits (SDKs) and GPU-enabled C programming language variants.

This paper will consider the parallelization of an anomaly detection algorithm based on the multivariate normal mixture model and the resulting parallel GPU implementations using CUDA. These implementations will be compared to an optimized multi-core CPU implementation, and processing performance will be evaluated for different parameters. Finally, by performing a simple anomaly detection experiment in a search and rescue scenario on a real pre-recorded hyperspectral image, it is shown that parallelization of the problem and the latest developments in GPU design have made real-time on-board normal mixture based anomaly detection feasible.

The outline of the paper is as follows: in Sect. 2 the anomaly detection algorithm is presented. Section 3 discusses the parallelization of parts of this algorithm, while Sect. 4 considers the resulting parallel implementations. Experimental results are discussed in Sect. 5 and the final conclusions are presented in Sect. 6.

2 Anomaly detection algorithm

The anomaly detection algorithm used here is based on a global multivariate normal mixture model representation of the background clutter, as discussed in [5]. The basic steps in this processing are:

Algorithm 1 Anomaly detection

Input: A hyperspectral image

1. **Estimation:** Estimate a background model by fitting a multivariate normal mixture model to a spatial subset of the image.
 2. **Evaluation:** Calculate a background probability value for all pixels in the image based on the estimated model.
 3. **Detection:** Detect anomalous pixels by thresholding on low background probability values.
 4. **Segmentation:** Merge fragmented pixel detections to objects by performing a binary morphological closing¹ in the thresholded image.
-

The first two steps are the key elements in this method and also by far the most time consuming. The last two steps are considered here as post-processing, and will only be performed when evaluating detection performance. Since it is reasonable to assume that the detection and segmentation steps give insignificant contributions to the overall

¹ Morphology is discussed in most image processing textbooks, e.g. Section 8.4 in [4].

processing time, only the time spent on performing estimation and evaluation are considered in the following experiments.

Hyperspectral sensors usually record the images line by line in a “pushbroom” scanning mode. The simplest way to employ the above anomaly detection algorithm in a real-time application is to process the continuously recorded data in blocks, similar to what is done in the ARCHER system [19]. Each newly recorded block may thus be sent off to processing, provided that processing of the previous block is finished. If the processing rate is faster than the sensor acquisition rate, this results in a small latency equal to the time it takes to record a block of data. The crucial factor in enabling a real-time implementation of this algorithm is therefore to ensure that the normal mixture estimation and evaluation steps are performed faster than the time it takes to record a block of data. The following section will give a detailed explanation of the estimation and evaluation steps.

2.1 Normal mixture model estimation and probability value calculation

A hyperspectral image can be considered as a set of pixel vectors $\mathbf{X} = \{\mathbf{x}_j \in \mathbf{R}^B, j = 1, 2, \dots, n\}$, where n is the number of image pixels and B is the number of spectral bands (see Fig. 1).

A multivariate normal mixture model is represented by the probability density function:

$$p(\mathbf{x}) = \sum_{c=1}^C \omega_c \phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \quad (1)$$

where $\omega_c \in [0, 1]$ is the mixing proportion (or weight) of component c with $\sum_{c=1}^C \omega_c = 1$, and $\phi(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate normal density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

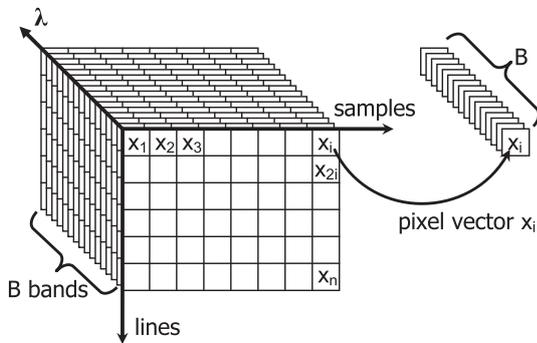


Fig. 1 Structure of the hyperspectral image data

$$\phi_c(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) = \frac{1}{(2\pi)^{B/2} |\boldsymbol{\Sigma}_c|^{1/2}} \times \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T \boldsymbol{\Sigma}_c^{-1} (\mathbf{x} - \boldsymbol{\mu}_c)\right\}. \quad (2)$$

Estimating a multivariate normal mixture model for the background is therefore equivalent to the problem of estimating the parameters $\boldsymbol{\psi} = \{C, \omega_c, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c; c = 1, 2, \dots, C\}$, given a set of image data. The total number of parameters that must be estimated is $P = (B(B+1)/2 + B+1)C + 1$ which in typical hyperspectral anomaly detection applications may be a quite large number. But since the background model estimation is based on data in the entire image block under consideration, more than enough data are available for the estimation process. In fact, the amount of data available may exceed that needed to make a statistically significant estimation of the model parameters. To avoid wasting time on processing more data than necessary, a subset of pixel vectors $\mathbf{S} = \{\mathbf{s}_j \in \mathbf{R}^B, j = 1, 2, \dots, m\}$, $\mathbf{S} \subseteq \mathbf{X}$, is considered where m is the number of pixels in the subset.

The actual estimation procedure used in this paper is an iterative method similar to the SEM algorithm [8], as outlined in Algorithm 2. The principal idea is to assume that each pixel \mathbf{s}_j from subset \mathbf{S} belongs to one of the components $c = 1, 2, \dots, C$. Thus, on each iteration i we obtain a partition $\mathbf{Q}_1^i, \mathbf{Q}_2^i, \dots, \mathbf{Q}_C^i$ of the subset \mathbf{S} , where $\mathbf{Q}_c^i = \{\mathbf{x}_j^i, c \in \mathbf{R}^B, j = 1, 2, \dots, m_c^i\}$ contains the pixels belonging to the component c on the iteration i , and m_c^i is the number of pixels in \mathbf{Q}_c^i .

Algorithm 2 Estimation step

Input:

- a subset \mathbf{S} of hyperspectral pixels
- an upper bound X_{max} on the number of components
- a threshold δ for termination of the iteration process
- a maximum number of iterations I_{max}

Output: The model parameters $\boldsymbol{\psi}$.

Initialization (Iteration 0):

Let $C = C_{max}$. Determine the first partition $\mathbf{Q}_c^0, c = 1, 2, \dots, C$ of \mathbf{S} :

1. Choose randomly C pixels from the subset \mathbf{S} to serve as component (cluster) centers (Fig. 2, Task 1).
2. Assign pixels of the subset \mathbf{S} to the components on the basis of the nearest Euclidean distance to the component center (Fig. 2, Task 2).

For every iteration $i > 0$ (I iterations in total):

Parameter estimation step:

Estimate $\boldsymbol{\mu}_c^i, \boldsymbol{\Sigma}_c^i$ and ω_c^i for $c = 1, 2, \dots, C$, using the component-wise Maximum Likelihood estimates (Fig. 2, Tasks 3-5):

$$\boldsymbol{\mu}_c^i = \frac{1}{m_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} \mathbf{x}_{j,c}^{i-1} \quad (3)$$

$$\boldsymbol{\Sigma}_c^i = \frac{1}{m_c^{i-1}} \sum_{j=1}^{m_c^{i-1}} (\mathbf{x}_{j,c}^{i-1} - \boldsymbol{\mu}_c^i)(\mathbf{x}_{j,c}^{i-1} - \boldsymbol{\mu}_c^i)^T \quad (4)$$

$$\omega_c^i = \frac{m_c^{i-1}}{m}. \quad (5)$$

Component assignment step:

1. Assign each pixel in the subset \mathbf{S} to one of the components (Fig. 2, Task 6) according to the maximum *a posteriori* probability criteria²:

$$\mathbf{x}_j \in \mathbf{Q}_c^i : \Pr(c|\mathbf{x}_j) = \max_l \Pr(l|\mathbf{x}_j) \quad (6)$$

where

$$\Pr(c|\mathbf{x}_j) = \frac{\omega_c^i \phi_c(\mathbf{x}_j; \boldsymbol{\mu}_c^i, \boldsymbol{\Sigma}_c^i)}{\sum_{c=1}^C \omega_c^i \phi_c(\mathbf{x}_j; \boldsymbol{\mu}_c^i, \boldsymbol{\Sigma}_c^i)}. \quad (7)$$

2. To avoid problems related to degenerations in the model, eliminate component c if $m_c^i < B$, $c = 1, 2, \dots, C$ (Fig. 2, Task 7). The pixels that belonged to the deleted components will be reassigned to the other components in the next iteration.

3. If the number of pixels from the subset \mathbf{S} that changed component membership is larger than the threshold δ and the number of iterations has not exceeded the maximum number of iterations I_{max} , return to the parameter estimation step.

Having estimated the multivariate normal mixture model for the background, a metric for correspondence with the background is calculated for each pixel in the hyperspectral image \mathbf{X} by evaluating the model probability density value for each pixel spectrum, as outlined in Algorithm 3 (see also Fig. 2, Task 8).

Algorithm 3 Evaluation Step

Input:

- a hyperspectral image \mathbf{X}
- a multivariate normal mixture model for the background $p(\mathbf{x}; \boldsymbol{\psi})$

Output: $p_{bg}(j)$

For each hyperspectral pixel \mathbf{x}_j in \mathbf{X} , calculate the background probability value

$$p_{bg}(j) = p(\mathbf{x}_j; \boldsymbol{\psi}) \quad (8)$$

² The original SEM algorithm uses the stochastic component assignment instead, a slower but more robust approach.

3 Parallelizing the anomaly detection algorithm

A block diagram of the anomaly detection algorithm is shown in Fig. 2. One of the characteristics of the algorithm is its regular (pipeline) structure. The figure gives the computational complexity for each algorithm task. We assume that the number of pixels in the original image block, as well as the subset used for the model estimation, is significantly larger than the number of components, number of bands and number of iterations in the estimation step ($n, m \gg C, B, I$). Then the overall computational complexity for the estimation step (Tasks 1–7) is $O(mCB^2I)$, and for the Evaluation step (Task 8) it is $O(nCB^2)$. Since in our case one block of hyperspectral data has spatial dimensions of order $10^3 \times 10^3$, the total number of pixels $n \sim 10^6$. We assume that $m \sim 10^5$ and $C, B \sim 10^1$. Generally, the computational cost of the anomaly detection algorithm is high. Thus, running the algorithm in real-time requires an efficient implementation and high-performance hardware.

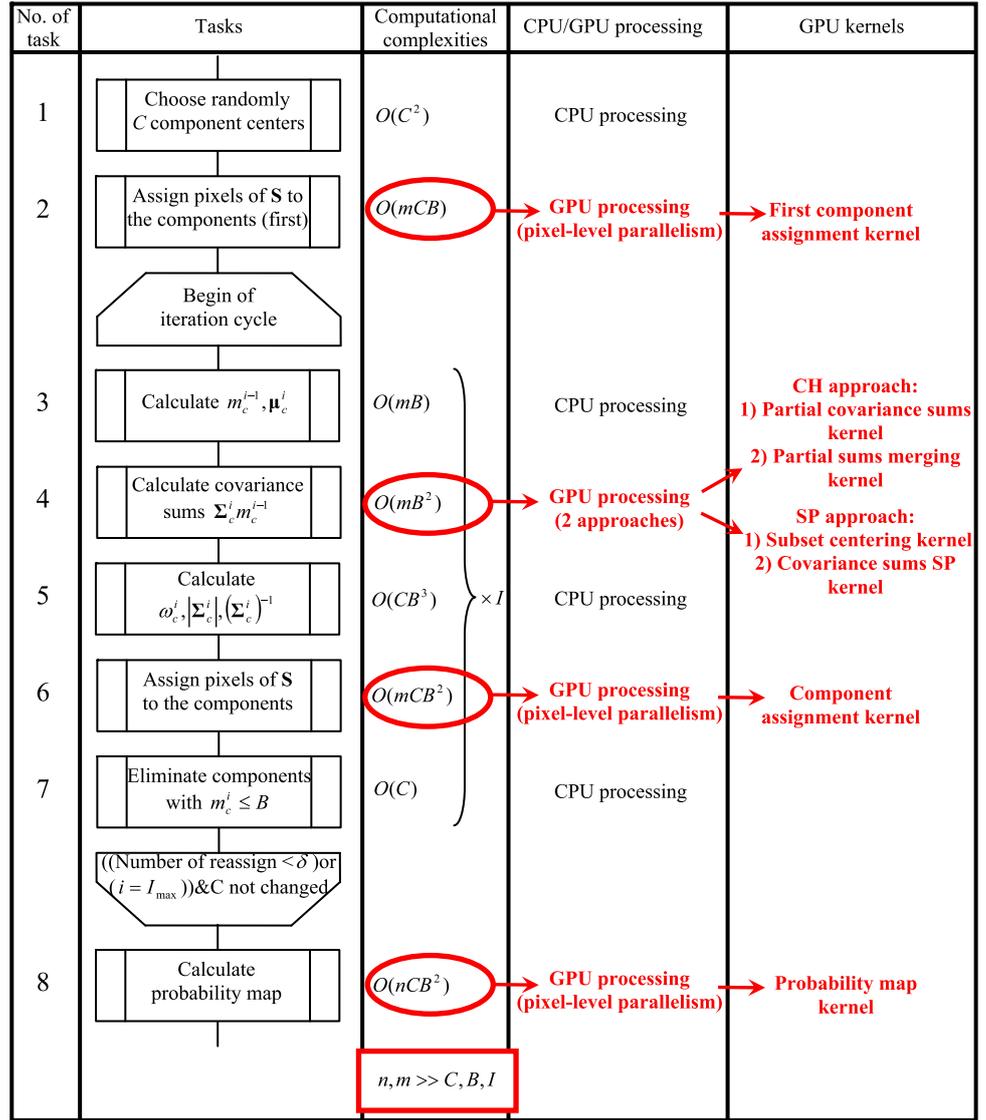
Plaza et al. [13] and Setoain et al. [15] have reviewed parallel processing of hyperspectral images. There are two main approaches to decompose the problem into parts that can be run concurrently: task-level decomposition and data-level decomposition [9, 11]. Setoain et al. [15] distinguish task-level, spatial-level and spectral-level parallelism for the hyperspectral image processing algorithms (the last 2 levels are the particular cases of the data decomposition patterns).

Task-level parallelism refers to different and independent sets of instructions executing in parallel. Spatial-level parallelism decomposes the image into subsets of pixel vectors that are operated on independently, thus forming data streams processed concurrently by the processing elements (the finest level being pixel-level decomposition, when each processing element is working on 1 pixel vector). Spectral-level parallelism refers to decomposition of the multi-band image data into units containing subsets of contiguous spectral bands.

Task-level parallelization is not possible here, as Fig. 2 shows that execution of each task requires the results from the previous task. Analysing the computational complexities of the parts of the algorithm, we can distinguish those with the highest computational cost as tasks 2, 4, 6, 8, marked by ellipses in Fig. 2. Fortunately, all these four tasks can be parallelized, using data-level decomposition.

The tasks that assign pixels to the components (Tasks 2, 6), and the evaluation task (Task 8) exhibit inherent parallelism at pixel level, the finest level of spatial parallelism. This results in simple, robust, scalable and easily understandable parallel implementation of these tasks. The number of threads that can be run concurrently is equal to the number of pixels (n, m). As the values of n, m are high, the amount of concurrency is significant. We note that all

Fig. 2 Block diagram of the anomaly detection algorithm (n number of image pixels, m number of pixels in the estimation subset \mathbf{S} , B number of bands, C number of components in the mixture, I number of iterations). *Red ellipses* indicate tasks with the highest computational cost. The diagram also summarizes the structure of GPU-based algorithm implementations, as discussed in the text



the concurrent threads of these tasks will require the common parameter data (like weights, means, inverse covariance matrices etc.). These data remain constant and can be efficiently shared between threads.

A more challenging step is the calculation of what we call the covariance sums in task 4: $\mathbf{Z}_c^i = m_c^{i-1} \Sigma_c^i$. Here $CB(B+1)/2$ elements must be estimated (symmetric covariance sum for each of the components). Several approaches to parallelize this task are possible. We consider two approaches.

3.1 Covariance sums: chunking approach (CH)

The first approach splits the hyperspectral image subset \mathbf{S} into K parts (chunks), and calculates the covariance sums for all the parts in parallel. Subsequently, covariance sums for the whole subset are calculated by summing in parallel the covariance sums for its parts (see Algorithm 4).

Figure 3b represents schematically these two branching steps.

Algorithm 4 Covariance sums - chunking approach

1. Decompose the subset S into K parts $U_k, k = 1, 2, \dots, K$ and execute in parallel $T = K$ threads:

for each thread k :

for each pixel $\mathbf{x}_{j,c}^{i-1} \in U_k$:

$$\mathbf{z}_{j,c}^{i-1} = \mathbf{x}_{j,c}^{i-1} - \mu_c^i$$

$$\mathbf{Z}_c^i = \mathbf{Z}_c^i + (\mathbf{z}_{j,c}^{i-1})(\mathbf{z}_{j,c}^{i-1})^T$$

end for

end for

2. Each element of the covariance sums for the subset S is calculated in parallel as the sum of the corresponding elements for the K chunks of the subset (in total $CB(B+1)/2$ elements, so $T = CB(B+1)/2$ threads are executed in parallel).

Regarding the scalability of the chunking approach, with the increase of the number of chunks K , more memory is needed to store intermediate covariance sums. Thus, there is an upper bound on K , and the scalability of the first step of the considered approach depends on the memory available and the memory bandwidth. The scalability of the second step is limited by the $CB(B + 1)/2$ concurrent threads. However, as the first step includes multiplication operations and in total more arithmetic operations per thread than the second step (for the typical configuration of values n , K and B), the complexity of the chunking approach is dominated by the first step.

3.2 Covariance sums: spectral-level parallelism (SP)

Another way to parallelize the covariance sums estimation is to calculate in parallel the covariance between bands q and r ($\mathbf{Z}(q, r)$, $q = 1, 2, \dots, B$; $r = 1, \dots, q$). Each thread will calculate C elements $\mathbf{Z}_c^i(q, r)$, $c = 1, 2, \dots, C$ (see Fig. 3c).

The algorithm consists of two branching steps: centering of the input subset S (in m parallel threads) and covariance sums calculation (see Algorithm 5).

Algorithm 5 Covariance sums - spectral-level parallelism

1. for each pixel $\mathbf{x}_{j,c}^{i-1} \in S$ ($T = m$ concurrent threads):

$$\mathbf{z}_{j,c}^{i-1} = \mathbf{x}_{j,c}^{i-1} - \boldsymbol{\mu}_c^i$$

end for

2. Each concurrent thread (in total $T = B(B + 1)/2$ threads) calculates $\mathbf{Z}_c^i(q, r)$, $c = 1, 2, \dots, C$.

The complexity of this algorithm is dominated by the second step, where $T = B(B + 1)/2$ threads are executed concurrently. As $B \sim 10^1$, the scalability here is seriously limited. This approach is interesting when the number of bands is significant.

In Sect. 5.3 below, we compare the execution speed and scalability of the two approaches for computation of covariance sums.

4 GPU-based parallel implementation

The previous section has shown that several tasks of the anomaly detection algorithm possess a significant amount of data-level concurrency, suitable for a “single instruction multiple data” architecture that allows massively parallel processing.

We have chosen to implement the parallel anomaly detection algorithms on an NVidia GeForce 8800 Ultra GPU, exploiting the new CUDA technology [12]. Through

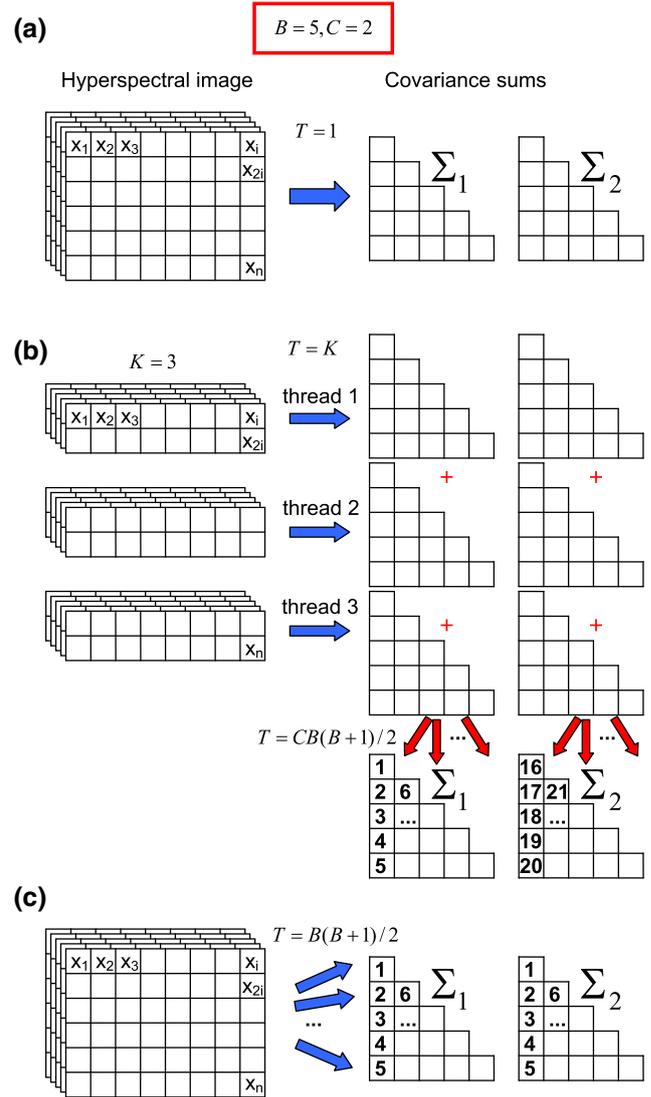


Fig. 3 Different approaches for calculation of the covariance sums: **a** sequential algorithm; **b** parallel algorithm—chunking approach; **c** parallel algorithm—spectral-level parallelism. The figure assumes $B = 5$, $C = 2$ and $K = 3$. The numbers in the covariance sums’ matrix cells correspond to different parallel threads

CUDA, the GPU (*device*) operates as a highly multi-threaded coprocessor to the main CPU (*host*). This means that the part of the program executed many times independently on different data can be isolated into a function (*kernel*), compiled to the device instruction set and executed concurrently on the device. The GPU is capable of running a very high number of threads in parallel.

The host and the device have their own DRAM (*host memory* and *device memory*, respectively). The data can be copied from one memory to another, by using the device’s high-performance Direct Memory Access engines. This improves significantly the data transmission performance, when compared to the previous GPU programming models.

Two main conditions must be fulfilled to achieve a good performance gain:

- Overlapping of memory accesses with arithmetic operations. GPU-based implementation is well suited for problems with high arithmetic intensities (ratio of arithmetic operations to memory accesses).
- Optimization of memory accesses. The device has on-chip shared memory (that threads can use for data sharing) with very fast read and write access and off-chip constant and texture cached memories. The high-bandwidth memory use must be maximized (like shared memory, cached accesses), while minimizing the accesses to uncached memory.

From the analysis in the previous section, the anomaly detection algorithm appears to fulfill these requirements reasonably well.

The most computationally demanding tasks of the algorithm have been implemented into seven GPU kernels as summarized in Fig. 2. A brief overview of the GPU/CUDA implementation for Tasks 2, 4, 6, 8 is given below:

- *Task 2—First component assignment kernel*: Each thread determines the normal mixture component with the minimal Euclidian distance between its center and the current pixel (each thread operates on one pixel), and stores the index of this component to the component membership array. Before executing the kernel, vectors of the C component centers are copied to the device constant memory. These values are cached once and afterwards they are used by each thread from the constant cache, thus optimizing the memory access time. In total $T = m$ threads are executed in this task.
- *Task 4*:

CH approach (refer Algorithm 4):

1. *Partial covariance sums kernel*: Each thread calculates the covariance sums (for C components) for one (current) chunk of the subset \mathbf{S} (in total $T = K$ threads), taking as input the component membership array and the means for normal mixture components. Before the kernel execution, the component means are mapped into the device texture memory (as a 2-dimensional CUDA array). These values are cached during the kernel execution. For each pixel, first a thread calculates its centered vector and store this vector to the shared memory. Then, this vector is used to calculate and add the contribution of the pixel to the covariance sum of the component, to which this pixel belongs. Each element of the vector will be read from the memory $B + 1$ times; therefore, the use of the

shared memory optimizes the memory access time.

2. *Partial sums merging kernel*: This kernel calculates covariance sums for C components, by summing the K partial covariance sums vectors, produced by the previous kernel. Each thread calculates one element of the covariance sums vector (which contains $CB(B + 1)/2$ elements). Thus, in total $T = CB(B + 1)/2$ threads are executed.

SP approach (refer Algorithm 5):

1. *Subset centering kernel*: Each thread calculates the centered vector for one pixel (in total $T = m$ threads), taking as input the means for normal mixture components (mapped into the device texture memory) and the component membership array.
2. *Covariance sums SP kernel*: Each thread calculates C elements $\mathbf{Z}_c^i(q, r)$, $c = 1, 2, \dots, C$ of the covariance sums vector (see Sect. 3 for details). The kernel takes as inputs the array of centered pixel vectors, produced by the previous kernel, and the component membership array. The elements $\mathbf{Z}^i(q, r)$ are kept in the shared memory during their calculation. In total $T = B(B + 1)/2$ threads are executed.

- *Task 6—Component assignment kernel*: Each thread operates on one pixel of the subset \mathbf{S} (in total $T = m$ threads), and assigns component membership according to (6). The kernel requires as inputs the parameters of the normal mixture model (weights, means and covariance matrices for C components). These parameters are stored in the device texture memory. The kernel's output is the component membership array. The intermediary vectors of centered pixel values (each vector is local for each thread) are kept in the local off-chip memory. They could be put in the shared memory as well, but as the size of the shared memory is limited (16 KB per multiprocessor for an NVidia GeForce 8800 Ultra), this will limit the number of threads running concurrently. Keeping these vectors in the local memory allows to run many threads in parallel, and the memory latencies (due to the access to the off-chip memory) are hidden by multithreading.
- *Task 8—Probability map kernel*: Each thread calculates for one pixel of the hyperspectral image \mathbf{X} a background probability value (1), in total $T = n$ threads. The parameters of the normal mixture model (weights, means and covariance matrices for C components) stored in the texture memory are used as inputs. The vectors of centered pixel values are kept in the local

off-chip memory (the same reasoning as for the *Component assignment kernel*). The resulting probability map is an important intermediate result of the anomaly detection algorithm.

The memory usage has been carefully optimized for all kernels, so that the fast shared memory and cached memories are used intensively. However, the device memory filling will depend on the size of the hyperspectral image \mathbf{X} , and the chosen subset size.

It can be noted that while CPU parallel code can be more easily adapted to different ranges of user parameters and data characteristics, the GPU code must ideally be designed for a specific problem size to have optimal performance. In our experiments, we use the same program for different ranges of parameters. Our code allows a range of reasonable parameters in the anomaly detection problem, but the performance may be sub-optimal for particular configurations of parameters.

5 Experimental results

5.1 Executing platforms and implementations

Our experiments were performed on a 2006-model HP xw8400 Workstation based on dual Quad-Core Intel Xeon processor E5345 running at 2.33 GHz with 1.333 MHz bus speed and 3 GB RAM. The computer was equipped with a XFX GeForce 8800 Ultra video card with 128 stream processors, 768 MB memory, 612 MHz core clock, 1,511 MHz shader clock and 2.16 GHz memory clock. This video card served as the primary display as well as a CUDA device.

Three different implementations of the anomaly detection algorithm have been made, one for the multi-core CPU and one GPU-based implementation for each of the covariance sum approaches (GPU-CH and GPU-SP). The CPU implementation is our performance reference, and also serves to check the precision and correctness of the GPU-based implementations.

Programs are built and run under the Windows XP 32-bit operating system. The CPU implementation is built with the Intel C++ Compiler 9.1 using OpenMP [3], BLAS [7] and LAPACK [2] libraries, while the GPU implementations has been made using the CUDA compiler driver `nvcc` [12] (CUDA Toolkit 1.0 and CUDA SDK 1.0 are used). For all implementations, the code has been carefully optimized including the mathematical representations, memory use and threading.

The dual Quad-Core Intel Xeon processor has eight cores, and therefore, up to eight threads can be executed in parallel on CPU. The parallel implementation on CPU is

efficient when a few concurrent threads execute relatively large number of operations (whereas GPU parallel implementations are efficient for executing a very high number of threads concurrently).

In our reference CPU-based implementation, Tasks 6 and 8 are implemented in parallel by means of OpenMP, so that each thread operates on one pixel (the same spatial-level parallelism as for the GPU-based implementations). As the anomaly detection algorithm includes a lot of operations on vectors, BLAS functions are used intensively throughout the program to optimize the processing time. Furthermore, the determinants and inverses of covariance matrices were computed using LAPACK functions. We also tried to run in parallel other parts of the program, but for the typical range of parameters in the anomaly detection problem the processing time was not reduced.

It can be also noted that the scalability of the CPU-based implementation is seriously limited by the number of processing cores available for the program execution. Currently, the number of CPU cores cannot be increased much beyond our eight-core desktop system before weight and power consumption becomes unacceptable for on-platform processing in many important cases such as airborne applications. Furthermore, the increase of performance through the generations of recent GPUs is faster than for CPUs.

5.2 Hyperspectral image data set

The hyperspectral data used here originate from a real airborne hyperspectral recording of a forest scene east of Oslo, Norway. The image was captured by a HySpex [10] visual and near infrared (VNIR) hyperspectral camera from an altitude of about 1,500 m above ground level. The HySpex VNIR module is a push-broom imager covering the spectral range from 0.4 to 1.0 μm in 160 spectral bands with 1,600 spatial pixels over a 17° cross-track field of view. The acquisition rate of the camera is about 100 lines/s or 0.16 Mpixels/s.

The 1,600 by 1,200 pixel (1.92 Mpixel) block used in the following experiments is extracted from the original hyperspectral image and is spectrally downsampled to 2–50 bands by averaging over neighbouring bands. In correspondence with several investigations into the number of bands required to obtain good target detection performance [6, 16], we expect to achieve good detection results in the lower half of this interval.

The targets used in the experiments are objects considered relevant in a search and rescue scenario. They are comprised of a green canvas textile similar to that one would find in some tents, and four sets of different coloured clothing laid on the ground in the direction of the four cardinal points north, east, south and west. The targets were

placed in plain view on a small marsh. Figure 4 shows photographs of the targets and the surrounding environment, while Table 1 describes each target in more detail.

5.3 Basic performance assessment

We evaluate the performance of the CPU and GPU-based implementations by measuring the execution time as a function of several parameters: the number of bands B , the number of components in the mixture C , the number of pixels m in the training set S and number of iterations I . Thus for our basic performance testing, the number of iterations is an input parameter and not controlled by a convergence criterion.

In the experiments we vary one parameter at a time, keeping the others fixed at the following standard configuration: $B = 15$ bands, $C = 10$ components, $I = 10$ iterations and a subset size of $m = 192,000$ pixels (10% of the whole image block). The execution time is measured for the complete execution as well as for individual parts. Here we report separately the contributions of the initialization part (Tasks 1–2) and the covariance matrix calculation part (Task 4) of the estimation step, and the time spent on the evaluation step (Task 8).

To determine the program execution time, the `C` function `clock()` was used for the CPU implementation and the CUDA timer was used to measure time for the GPU implementations. The total time measurement is started right after the hyperspectral image file is read to the CPU memory and stopped right after the resulting probability map is obtained and stored in the CPU memory. For timing of the individual parts, memory transfers related to these parts are included.

The measurements were found to be repeatable within about 1% for the GPU implementations. For the CPU implementation the variation was somewhat larger, probably due to interrupts and task scheduling by the operating system, although there is still good consistency across the explored range of parameters. For real-time applications it

is interesting to note that the GPU execution time measurements are very stable. This means that the GPU may be run closer to its peak performance, with less needs for time margins compared to the CPU.

Figure 5 shows the measured total execution time when varying the different parameters. Not surprisingly, the execution time scales approximately linearly with the number of components C , iterations I , and subset size m . With increasing band count B , the increase in execution time is somewhat faster than linear. The overall result is that the GPU increases computing speed by a significant factor. The gain is particularly large for lower band counts, for example more than 20 times faster for 5 bands. At 15 bands the speedup is a factor 10, while at 50 bands a more modest factor of 3 is obtained.

The lower gain at high band count is essentially due to the covariance sums computation which becomes more memory intensive and hence less adapted to GPU processing for increasing covariance matrix dimensionality. As Fig. 6 shows, the CPU implementation performs comparable to or better than the GPU-CH implementations for most band counts during the covariance sum processing, while the GPU-SP is much slower than the other implementations below 25 bands.

Analysing the algorithms of covariance sums computation, several reasons can be suggested why the GPU-CH implementation for this task is slower than the CPU-based one. For a small number of bands the calculation time is spent mostly to run through all the array of pixel vectors.

Table 1 Target descriptions

Name	Description
A	Green canvas, about 1.5×2.5 m
B	Jeans jacket and pants
C	Grey coat
D	Red jacket and pants
E	Green jacket and pants

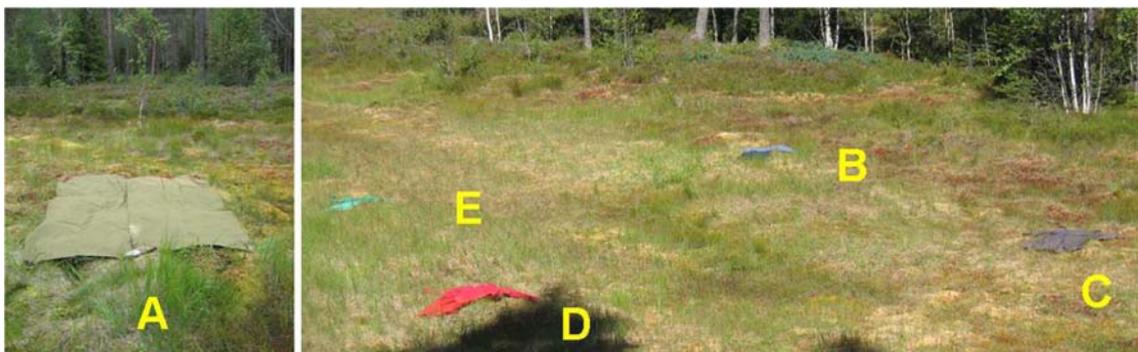


Fig. 4 Target layout and the names used to refer to them in the following experiments. See Table 1 for more information

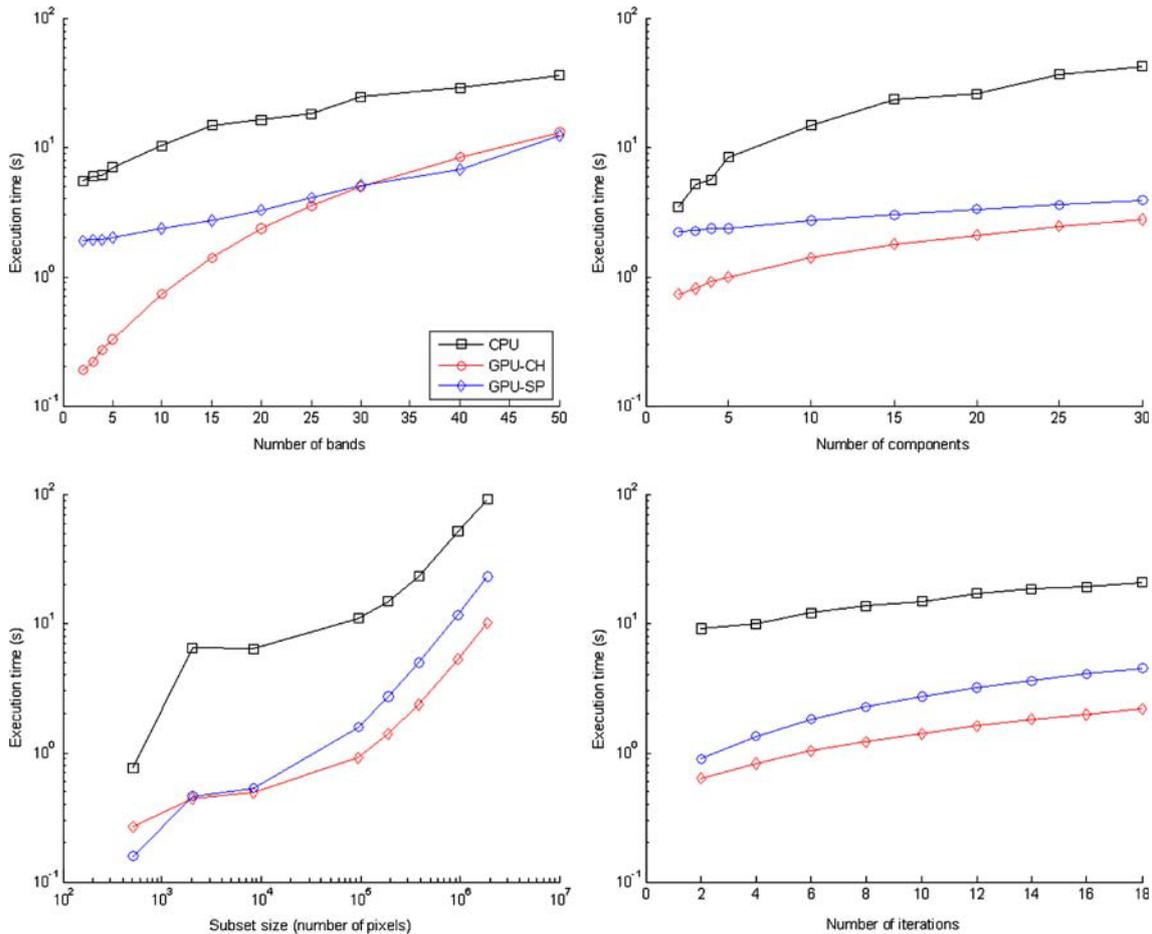


Fig. 5 Total execution time for the three implementations when varying different parameters. The plots show the execution time for different choices of number of bands (*top left*), number of components

(*top right*), subset size (*bottom left*) and number of iterations (*bottom right*). Here, the default configuration used is 15 bands, 10 components, 10 iterations and a subset size of 192,000 pixels

When we split this array of pixel vectors into several parts (chunks) in the CH approach, the GPU execution time for this parallel approach becomes faster. But when the number of bands increases, the running through bands becomes more computationally demanding. In this case:

1. More memory is needed to store covariance sums for K chunks. As they are stored in the device global memory, memory bandwidth causes the increase of the processing time, when compared to CPU implementation. The processing on CPU allows data caching, which becomes especially advantageous when the number of bands increases.
2. As was mentioned before, the GPU code must be designed for a specific problem size and thread configuration to have optimal performance. A GPU kernel is executed in parallel by the batch of threads, organized as a grid of thread blocks [12]. The number of blocks and threads per block must be chosen to maximize performance. Furthermore, for the CH approach of covariance sums computation the number

of chunks K must be chosen. The GPU code was optimized for the standard configuration of parameters ($B = 15$ bands, $C = 10$ components, $I = 10$ iterations and $m = 192,000$ pixels). In particular, the number of chunks $K = 512$ was chosen by the experimental tuning and fixed in the program. As can be seen from Fig. 6, the GPU-CH implementation is the fastest for this configuration of parameters (when $B = 15$ bands, the processing time for the GPU-CH implementation is 570 versus 720 ms for the CPU implementation). If the GPU-code is adapted for another configuration of parameters, the processing speed may be increased for this particular configuration.

3. It must be noted that we varied the number of bands B , while keeping the estimation subset size m constant. However, with increasing B , the number of parameters of the multivariate normal mixture model increases, and larger subset of pixels is needed to obtain an accurate estimate of parameters. When varying the subset size m together with the number of bands B , the

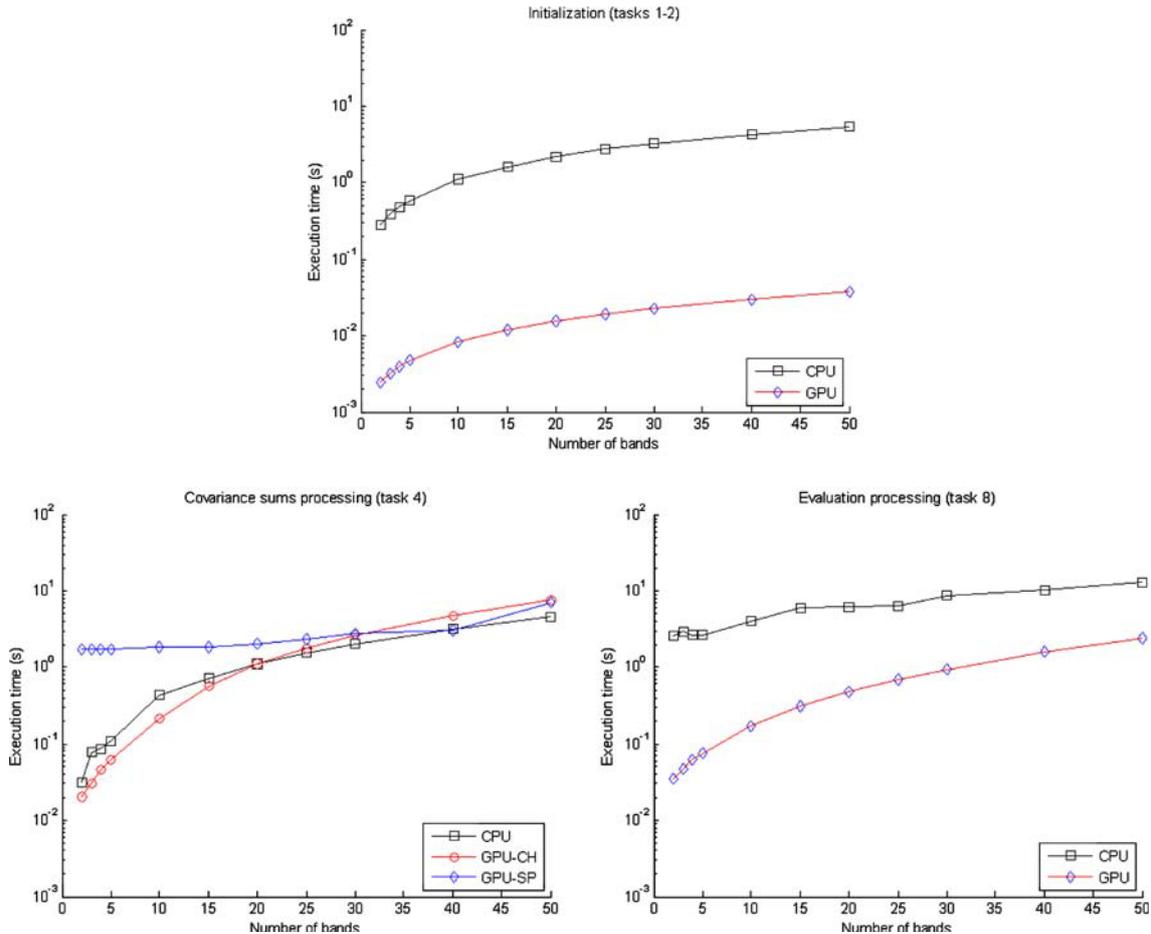


Fig. 6 Execution time for important parts of the implementations. The plots show the difference between the implementations in execution time for the initialization tasks (top), the covariance sums processing task (bottom left) and the evaluation processing task

(bottom right). Here, the bands are varied while keeping the other parameters fixed at 10 components, 10 iterations and a subset size of 192,000 pixels

GPU-CH implementation is likely to become more efficient, relative to CPU, for higher number of bands.

The GPU-SP implementation becomes interesting when the number of bands $B > 25$. The reason can be deduced from the algorithm, which explores a spectral-level parallelism. The GPU-SP implementation is faster than both other implementations when $B = 40$. However, for $B = 50$ it is slower than the CPU implementation. The probable reason is that the GPU-code is not well tuned for this problem size.

It is also evident that the gain in GPU-based evaluation processing is decreasing with higher band counts, although it is still significant for 50 bands. Interestingly, the GPU implementations of the initialization part achieves a speed-up gain of around 100. Since most of the initialization corresponds to significant parts of the K-means clustering algorithm, this result also demonstrates that parallel implementations of K-means on GPUs can give a significant increase in computing speed.

5.4 Real-time anomaly detection demonstration

After establishing that the parallel GPU implementations are significantly faster than the CPU implementation, we will now demonstrate the impact this has on anomaly detection processing. This experiment will consider the anomaly detector described in Sect. 2 applied in a search and rescue context. While a typical application would process the data in several consecutive blocks, we will here consider the processing of only one such block, and assume that the results obtained are representative for a string of blocks in average over time. Real-time performance is evaluated by comparing the block processing time with the actual time it took to record the block with the hyper-spectral camera.

As opposed to the previous experiment, the iterative procedure involved in the estimation process will here stop only when the convergence criterion is satisfied. For this demonstration a convergence threshold of $\delta = 3\%$ was

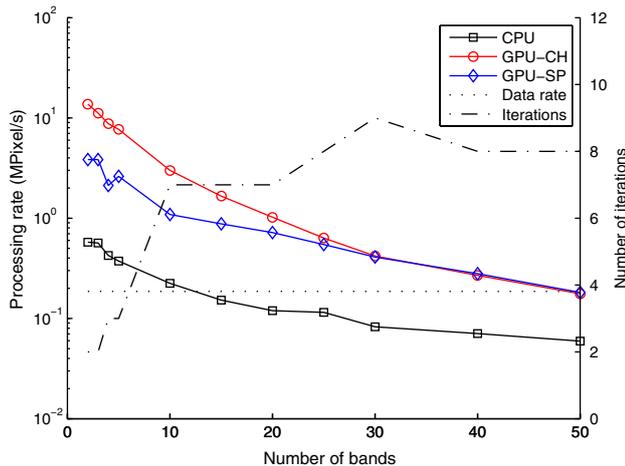


Fig. 7 Anomaly detection processing rate for different choices of number of bands. The *solid lines* show the processing rate for the different implementations, while the *dotted line* shows the sensor data rate (see the left y-axis). The *dash-dot line* shows the number of iterations needed to reach convergence for the different band counts (see the right y-axis)

chosen. In addition, the number of components was chosen to be $C_{\max} = 10$ and the size of the estimation subset was set to be 10% of the pixels in the image block.

Figure 7 shows the processing rates when each of the three implementations is applied to the different spectrally downsampled images. To be fairly certain that the observed rates are not an extreme result from the random initialization of the estimation, the median rate of 19 runs is chosen for each implementation and band configuration. By comparing with the sensor data rate, represented by the dotted line in Fig. 7, we see that the parallel GPU-based implementations run faster than the data rate right up to about 50 bands. Hence, by exploiting the power of GPU processing, multivariate normal mixture based anomaly detection can be run in real time under similar conditions for less than 50 spectral bands on current hardware. In the 15–25 band interval, the GPU implementations are about 3–10 times faster than the real-time constraint, while the CPU implementation is slower than real time above 10 bands.

When comparing Fig. 7 with Fig. 5 it is clear that the implementations are somewhat faster in this experiment. This is simply because the estimation process needs fewer iterations before satisfying the convergence criterion. Figure 7 also shows the number of iterations needed for the different band configurations.

To fully justify the claim that multivariate normal mixture based anomaly detection is performed in real time, sufficiently good detection results must be demonstrated. The detection results for the GPU-CH implementations are presented in Fig. 8. For 20 bands all the targets are detected with less than 1 false alarm per s, and 3 targets are detected

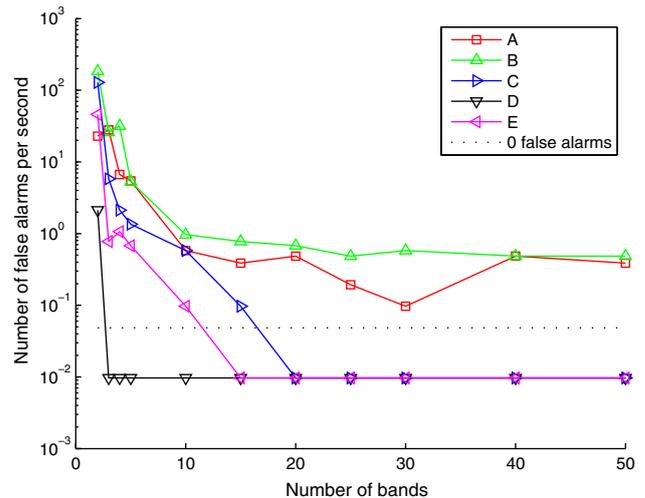


Fig. 8 Detection results for the GPU-CH approach for different choices of number of bands. Detection results without false alarms are placed below the *dotted line*

without false alarms. These are considered acceptable results for the target detection scenario in question, and may be further improved by exploiting the available processing time to use more accurate model estimation techniques and perform different false alarm mitigation methods (e.g. [14]). Figure 9 shows the detection result for 20 bands with a detection threshold set so that all targets are detected.

6 Conclusion

Multivariate normal mixture models form the basis of an algorithm for anomaly detection in hyperspectral images. The algorithm possesses a significant amount of data-level concurrency in its time-consuming parts, and appears well adapted to the GPU architecture. We have used CUDA to implement the computationally intensive parts of the algorithm on an Nvidia GeForce 8800 GPU, and compared its performance to a CPU-based implementation running on a dual quad-core computer.

Generally, the GPU provides a significant speedup of the algorithm compared to the CPU implementation. The relative performance of the GPU depends on the algorithm parameters such as data size and band count. Furthermore, it is often difficult to optimize GPU code without adapting it to a narrow range of parameters. For the pixel-parallel parts of the algorithm, speedups on the order of 10 and even 100 are observed. For the computation of covariances, however, the GPU only provides an advantage over the CPU for band counts below about 20. For higher band counts, the memory model of the GPU

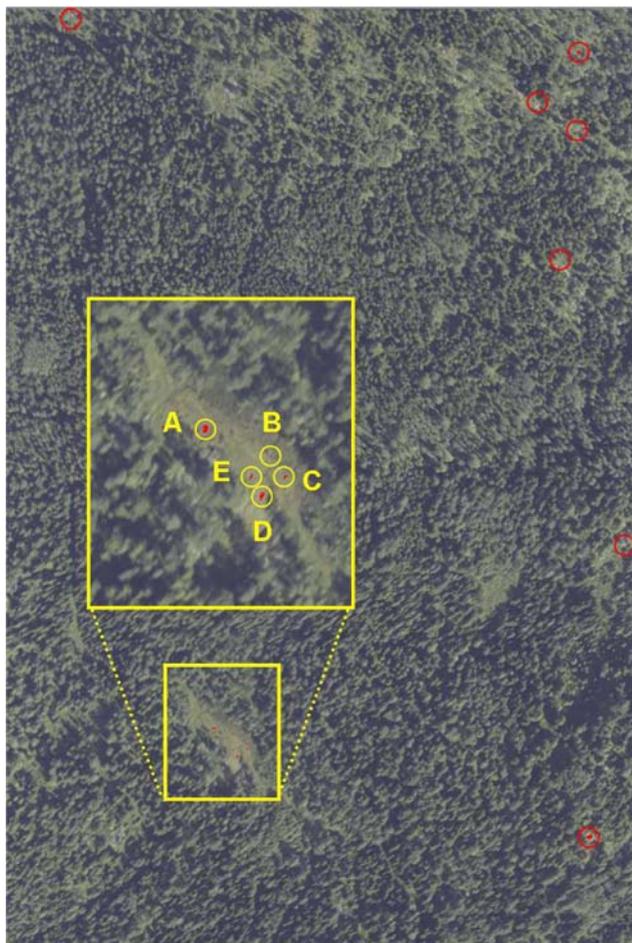


Fig. 9 Detection results for the GPU-CH approach when detecting all targets using 20 bands. The *red circles* show the false alarms, while the *yellow circles* show the detected targets. The image in the background is a colour image extracted from the hyperspectral image block used in the experiments

does not provide a speed advantage in the calculation of covariance sums.

Anomaly detection has been performed on a realistic hyperspectral data set. We have shown, crucially, that the GPU enables real-time execution of the algorithm on a hyperspectral data stream with high spatial and spectral resolution, with acceptable detection performance and a significant margin on computing time. This margin enables the same hardware to execute other parts of the detection system such as threshold estimation, spatial analysis, false alarm mitigation or signature-based spectral detection.

Finally, it can be noted that methods based on multivariate normal mixtures are versatile statistical tools with potential use in many areas beyond remote sensing. Up to now, computational complexity has precluded their use in many applications. This is about to become history with

the advent of highly parallel processing in desktop computers.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. AMD: ATI CTM Guide (2006)
2. Anderson, E., Bai, Z., Bischof, C., Blackford, L.S., Demmel, J., Dongarra, J.J., Croz, J.D., Hammarling, S., Greenbaum, A., McKenney, A., Sorensen, D.: LAPACK Users' Guide (3rd edn). Society for Industrial and Applied Mathematics, Philadelphia (1999)
3. Chandra, R., Menon, R., Dagum, L., Kohr, D., Maydan, D., McDonald, J.: Parallel Programming in OpenMP. Morgan Kaufmann Publishers Inc., San Francisco (2001)
4. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Addison-Wesley Publishing Company, Reading (1993)
5. Kåsen, I., Goa, P.E., Skauli, T.: Target detection in hyperspectral images based on multicomponent statistical models for representation of background clutter. In: Proceedings of the SPIE. 5612, pp. 258–264 (2004)
6. Kåsen, I., Rødningsby, A., Haavardsholm, T.V., Skauli, T.: Band selection for hyperspectral target detection based on a multinormal mixture anomaly detection algorithm. In: Proceedings of the SPIE. 6966, p. 696606 (2008)
7. Lawson, C.L., Hanson, R.J., Kincaid, D.R., Krogh, F.T.: Basic linear algebra subprograms for fortran usage. ACM Trans. Math. Softw. **5**(3), 308–323 (2006)
8. Masson, P., Pieczynski, W.: SEM algorithm and unsupervised statistical segmentation of satellite images. IEEE Trans. Geos. and Remote Sens. **31**(3), 618–633 (1993)
9. Mattson, T.G., Sanders, B.A., Massingill, B.L.: Patterns for Parallel Programming. Addison-Wesley Professional, Boston (2005)
10. Norsk Elektro Optikk AS.: For further information about the HySpex camera: <http://www.neo.no/hypspec/> (2008)
11. Nicolescu, C., Jonker, P.: A data and task parallel image processing environment. Parallel Comput. **28**(7–8), 945–965 (2002)
12. NVIDIA: NVIDIA CUDA compute unified device architecture—programming guide. <http://developer.nvidia.com/cuda> (2007)
13. Plaza, A., Valencia, D., Plaza, J., Martinez, P.: Commodity cluster-based parallel processing of hyperspectral imagery. J. Parallel Distrib. Comput. **66**(3), 345–358 (2006)
14. Schaum, A.: A remedy for nonstationarity in background transition regions for real time hyperspectral detection. In: 2006 IEEE Aerospace Conference, p. 9 (2006)
15. Setoain, J., Tenllado, C., Prieto, M., Valencia, D., Plaza, A., Plaza J.: Parallel hyperspectral image processing on commodity graphics hardware. In: Proceedings of the ICPPW '06, pp. 465–472 (2006)
16. Shen, S.S., Bassett, E.M.: Information-theory-based band selection and utility evaluation for reflective spectral systems. In: Proceedings of the SPIE. 4725, pp. 18–29 (2002)
17. Stein, D.W.J., Beaven, S.G., Hoff, L.E., Winter, E.M., Schaum, A.P., Stocker, A.D.: Anomaly detection from hyperspectral imagery. IEEE Signal Process Mag. **19**(1), 58–69 (2002)
18. Stellman, C.M., Olchowski, F.M., Michalowicz, J.V.: WAR HORSE (Wide Area Reconnaissance—Hyperspectral Overhead Real-time Surveillance Experiment). In: Proceedings of the SPIE. **4379**, pp. 339–346 (2001)

19. Stevenson, B., O'Connor, R., Kendall, W., Stocker, A., Schaff, W., Alexa, D., Salvador, J., Eismann, M., Barnard, K., Kershensstein, J.: Design and performance of the Civil Air Patrol ARCHER hyperspectral processing system. In: Proceedings of the SPIE. **5806**, pp. 731–742 (2005)

Author Biographies

Yuliya Tarabalka has received her Bachelor degree in Computer Science (2005) from the Ternopil Ivan Pul'uj State Technical University, Ukraine and M.Sc. degree in Signal and Image Processing (2007) from National Polytechnic Institute of Grenoble (INPG), France. From July 2007 to January 2008 she worked as a researcher with the Norwegian Defence Research Establishment, Norway. Since 2007, she is a Ph.D. student, pursuing a co-joint degree between University of Iceland, Iceland and INPG, France. Her current research work is funded by the "HYPER-I-NET" Marie Curie Research Training Network. Her research interests are in the areas of image processing, pattern recognition, hyperspectral imaging and development of efficient algorithms.

Trym Vegard Haavardsholm received his M.Sc. in Modelling and Data analysis at the University of Oslo, Norway in 2005. He is currently working at the Norwegian Defence Research Establishment as a scientist and Ph.D. student. His main research interests include image analysis and statistical modelling in hyperspectral and lidar images.

Ingebjørg Kåsen holds a M.Sc. in mathematics from the University of Oslo, Norway. She has worked as a research scientist at FFI, the Norwegian Defence Research Establishment, since 2002. Research interests include hyperspectral imaging, signal and image processing and statistical modelling.

Torbjørn Skauli holds a Ph.D. in physics from the University of Oslo and has worked at FFI, the Norwegian Defence Research Establishment, since 1990. Skauli has worked with semiconductor physics, infrared detectors, electronics, nonlinear optics and nanotechnology. Since 2001 he has been head of the hyperspectral imaging group at FFI.

