

**A SUPERVISED CLASSIFICATION SYSTEM OF SPECTRAL IMAGES ACQUIRED
WITH A SINGLE PIXEL OPTICAL ARCHITECTURE AND SIDE INFORMATION**

KAREN YANETH SÁNCHEZ QUIROGA

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE
TELECOMUNICACIONES
MAESTRÍA EN INGENIERÍA ELECTRÓNICA
BUCARAMANGA, 2019

**A SUPERVISED CLASSIFICATION SYSTEM OF SPECTRAL IMAGES ACQUIRED
WITH A SINGLE PIXEL OPTICAL ARCHITECTURE AND SIDE INFORMATION**

Author:
KAREN YANETH SÁNCHEZ QUIROGA

Research work to qualify for the Master degree in Electronic Engineering

Director:
HENRY ARGUELLO FUENTES
Ph.D. in Electrical and Computer Engineering

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE
TELECOMUNICACIONES
MAESTRÍA EN INGENIERÍA ELECTRÓNICA
BUCARAMANGA, 2019

A Dios.

A mi madre, Yaneth Quiroga, por el don de la vida, por darme la fuerza para continuar cada día, por su gran ejemplo de nobleza, amor infinito y entrega incondicional.

A mi padre, Carlos Enrique Sánchez, por ser el mejor en todo lo que hace y enseñarme con su ejemplo que para ser el mejor del mundo solo hace falta tomar la decisión.

A mis hermanos, Karol S. y Carlos E.

Karen Yaneth

AGRADECIMIENTOS

La autora expresa su agradecimiento:

Al profesor Henry Arguello Fuentes, director de este trabajo de investigación, por ser mi mentor, brindarme su apoyo, orientación y acompañamiento a lo largo de esta etapa de mi formación.

A Colciencias, la Universidad Industrial de Santander, y el Grupo de investigación HDSP por otorgarme los recursos necesarios para la ejecución de este proyecto.

A los integrantes del Grupo de investigación HDSP por sus enseñanzas, acogida y colaboración durante esta etapa de aprendizaje.

CONTENTS

INTRODUCTION	12
1 THEORETICAL FRAMEWORK	16
1.1 SUPERVISED CLASSIFICATION OF SPECTRAL IMAGING	16
1.1.1 Spectral Imaging.	16
1.1.2 Supervised Classification.	18
1.2 COMPRESSIVE SPECTRAL IMAGING	19
1.3 MULTI-RESOLUTION IMAGE FUSION	24
2 SUPERVISED SPATIO-SPECTRAL CLASSIFICATION OF FUSED IMAGES USING SUPERPIXELS	26
2.1 MATRIX MODEL	26
2.2 PROPOSED SPATIO-SPECTRAL CLASSIFICATION APPROACH	27
2.3 SIMULATIONS AND RESULTS	30
3 SUPERVISED CLASSIFICATION OF SPECTRAL IMAGES ACQUIRED WITH SINGLE PIXEL CAMERA AND SIDE INFORMATION	38
3.1 MATHEMATICAL SENSING MODEL	38
3.1.1 Side information-based SPC sensing model for spectral imaging.	38
3.1.2 Coded aperture design based on Super-pixel.	40
3.1.3 Reconstruction with Side Information.	42
3.1.4 Fast Multi-Resolution (FMR) Reconstruction.	42
3.2 SUPERVISED CLASSIFICATION MODEL FOR SPC MEASUREMENTS AND SIDE INFORMATION	43
4 EXPERIMENTS AND RESULTS	45
4.1 SIMULATIONS	45
4.2 EXPERIMENTAL SETUP	52
4.3 ACQUIRED DATASETS	53
4.4 ANALYSIS OF EXPERIMENTAL RESULTS	55
5 CONCLUSIONS	59
BIBLIOGRAPHY	60
APPENDIX	66

LIST OF FIGURES

1	Different types of multiband images according to their resolution	16
2	Traditional snapshot spectral imaging technologies	17
3	Spectral imaging concept	17
4	Optical architecture of the SPHIS	21
5	Single pixel combined with spectrometer.	22
6	Pansharpening	25
7	Flowchart of the proposed classification framework	27
8	Datasets: False-color, ground truth maps and spectral signatures	31
9	Overall accuracy of classification varying the number of superpixels	32
10	Overall accuracy of classification varying the number of training samples	33
11	Visual classification results on Pavia University image	35
12	Visual classification results on Pavia Center image	36
13	Visual classification results on Salinas Valley image	37
14	Scheme of the SPC used to acquire the spatial and spectral information with RGB side information.	39
15	SPC sensing matrix $\hat{\mathbf{H}}$, with $M = 2$, $N = 5$, $\gamma = 0.5$ and $L = 4$	40
16	RGB sensing matrix $\hat{\mathbf{H}}$, for $M = 2$, $N = 5$, and $L = 9$	40
17	Scheme of the process to generate MR coded apertures	41
18	Examples of the designed MR coded apertures	42
19	False-color composite of the datasets used	45
20	Diagram of tested approaches	46
21	Overall accuracy of classification varying the number of superpixels	47
22	Visual classification results on Pavia University dataset	49
23	Visual classification results on Pavia Center dataset	50
24	Visual classification results on Salinas Valley dataset	51
25	Testbed implemented to acquire the SPC and RGB measurements	52
26	Top view of the implemented architecture.	52
27	Process of the cob to the fermented cocoa beans.	53
28	Traditional classification of cocoa bean in Colombia	54
29	Cocoa beans of each fermentation category	55
30	False-color composite of mixed cocoa beans	55
31	Overall accuracy of classification varying the number of superpixels	56
32	Visual classification results on Cocoa bean 1 dataset	57
33	Visual classification results on the Cocoa bean 2 dataset	58

LIST OF TABLES

1 Quantitative results of the different classification methods for the Pavia University image (OA%).	34
2 Quantitative results of the different classification methods for the Pavia Center image (OA%).	34
3 Quantitative results of the different classification methods for the Salinas Valley image (OA%).	36
4 Quantitative results on Pavia University dataset (OA %)	49
5 Quantitative results on Pavia Center dataset (OA %)	50
6 Quantitative results on Salinas Valley dataset (OA %)	51
7 Quantitative results on Cacao bean 1 dataset (OA %)	57
8 Quantitative results on Cacao bean 2 dataset (OA %)	58

Resumen

TITULO: UN SISTEMA DE CLASIFICACIÓN SUPERVISADO DE IMÁGENES ESPECTRALES ADQUIRIDAS CON ARQUITECTURA ÓPTICA DE ÚNICO PIXEL E INFORMACIÓN COMPLEMENTARIA. *

AUTOR: KAREN YANETH SÁNCHEZ QUIROGA **

PALABRAS CLAVE: Imágenes espectrales, resolución espacial, muestreo compresivo, sentido espectral de imágenes, fusión de imágenes, clasificación supervisada, superpíxeles.

Las imágenes espectrales proporcionan una gran cantidad de información que permite realizar diversas tareas de procesamiento, como clasificación, con gran precisión. Sin embargo, debido a la alta dimensionalidad de los datos, procesar, transmitir y almacenar dicha información es costoso. En los últimos años, la compresión de imágenes espectrales (CSI) ha emergido como una nueva técnica de adquisición que adquiere proyecciones codificadas de la escena espectral aplicando diferentes patrones de codificación, reduciendo considerablemente los costos de almacenamiento y transmisión. Variando la estrategia de muestreo, varios dispositivos CSI, con diferentes configuraciones ópticas, se han desarrollado, donde la arquitectura de cámara de un solo píxel (SPC) sobresale por bajo costo de implementación. Tradicionalmente, una reconstrucción completa de la escena subyacente es necesaria para realizar cualquier tarea de procesamiento, lo que implica resolver un problema de optimización computacionalmente costoso. El objetivo de este proyecto es realizar clasificación de imágenes espectrales utilizando directamente mediciones CSI, evitando la reconstrucción completa de la escena. Las mediciones de CSI se adquirirán mediante la implementación en el laboratorio de una SPC. Además, dada la baja resolución espacial del sensor SPC, se propone obtener información complementaria a través de un sensor RGB auxiliar, que tiene una resolución espacial más alta. Utilizando la información de ambos sensores, este trabajo propone diseñar los patrones de codificación SPC considerando la agrupación de píxeles con características similares en la imagen RGB. Luego, es posible extraer características de la escena para realizar una clasificación directa. Por lo tanto, es posible obtener un mapa de clasificación, utilizando una máquina de soporte vectorial de manera rápida y con alta precisión sin requerir una etapa de reconstrucción. En general, se obtuvo una precisión global de 95.41%, 97.29%, 97.72% y 99% utilizando la "Pavia University", "Pavia Center", "Salinas", y "granos de cacao" adquiridos en un laboratorio óptico, respectivamente.

* Trabajo de Investigación.

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director, Henry Arguello Fuentes.

Abstract

TITLE: A SUPERVISED CLASSIFICATION SYSTEM OF SPECTRAL IMAGES ACQUIRED WITH A SINGLE PIXEL OPTICAL ARCHITECTURE AND SIDE INFORMATION. *

AUTHOR: KAREN YANETH SANCHEZ QUIROGA **

KEYWORDS: Spectral imaging, spatial resolution, compressive sensing, compressive spectral imaging, image fusion, supervised classification, superpixels.

Spectral images provide a large amount of information allowing to perform different processing tasks, such as classification, with greater precision. However, due to the data high-dimensionality, processing, transmitting, and storing such information is expensive. In recent years, Compressive Spectral Imaging (CSI) has emerged as a new acquisition technique which acquires coded projections of the spectral scene by applying different coding patterns, considerably reducing storage and transmission costs. Varying the sampling strategy, several CSI devices, with different optical configurations, have been developed, where the single pixel camera architecture (SPC) excels with a low implementation cost. Traditionally, a complete reconstruction of the underlying scene is needed to perform any processing task, which involves solving a computationally expensive optimization problem. The objective of this project is to perform spectral image classification by directly using the CSI measurements, thus avoiding the complete reconstruction of the scene. The CSI measurements will be acquired through the implementation in the laboratory of a SPC. Besides, given the low spatial resolution of the SPC sensor, it is proposed to obtain side information through an auxiliary RGB sensor, which has a higher spatial resolution. Using the information from both sensors, this work proposes to design the SPC coding patterns considering the grouping of pixels with similar characteristics in the RGB image. Then, it is possible to extract valuable features of the scene to perform a direct classification. Hence, it is possible to obtain a classification map, using a support vector machine, in a fast way and with high accuracy in the classification without requiring a reconstruction stage. In general, a global precision of 95.41%, 97.29%, 97.72% and 99% was obtained using the "Pavia University", "Pavia Center", "Salinas", and spectral images of "cocoa beans" acquired in an optical laboratory, respectively.

* Research Work.

** School of Physical-Mechanical Engineering. Department of Electrical, Electronics and Telecommunications Engineering. Advisor, Henry Arguello Fuentes.

INTRODUCTION

Spectral imaging (SI) is a technique used to acquire spatial information of a scene in many contiguous electromagnetic frequencies. The acquired image is commonly represented as a three-dimensional data cube, where two dimensions correspond to the spatial information and the third one to the spectral domain.¹ In general, SI is a valuable tool for multiple applications, including precision agriculture, medical images, food quality control, military surveillance, etc.² Based on the spectral/spatial resolution, spectral range, and width and contiguousness of bands, SI sensors can be categorized as multispectral (MS) or hyperspectral (HS). MS sensors can acquire up to 10 spectral bands. HS devices, on the other hand, capture up to hundreds of spectral bands of the scene of interest in a wavelength interval ranging from the visible region ($0.4 - 0.7 \mu\text{m}$) to the infrared region ($\approx 2.4 \mu\text{m}$).³

However, although HS sensors can provide abundant spectral information, their spatial resolution is generally low, mainly due to an unsurpassable tradeoff between spatial and spectral sensitivities⁴ as well as the atmospheric scattering, secondary illumination, changing viewing angles, sensor noise, etc.

On the other hand, due to the scanning patterns of conventional methods of spectral image acquisition, the amount of information captured is abundant. These methods, such as whiskbroom or pushbroom, are based on the Nyquist-Shannon theory⁵ and are shown in Section 1. Accordingly, in addition to the low spatial resolution, the high storage cost is a limitation of the HS images.

As a consequence of the low spatial resolution of this type of data, the problem of merging a high spectral and low spatial resolution image with an auxiliary image of higher spatial resolution but lower spectral resolution, also known as multi-resolution image fusion, has been explored for many years and is still an active, challenging and innovative field of research.⁶ Based on the used theoretical approach, data fusion algorithms can be categorized as component substitution, spectral unmixing, sparse signal representation, and Bayesian estimation. Traditionally, these algorithms fuse the information acquired

¹ José M Bioucas-Dias et al. "Hyperspectral remote sensing data analysis and future challenges". In: *IEEE Geoscience and remote sensing magazine* 1.2 (2013), pp. 6–36.

² D Lorente et al. "Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment". In: *Food and Bioprocess Technology* 5.4 (2012), pp. 1121–1142.

³ Gabriel Martín and José M Bioucas-Dias. "Hyperspectral blind reconstruction from random spectral projections". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.6 (2016), pp. 2390–2399.

⁴ Nasser M Nasrabadi. "Hyperspectral target detection: An overview of current and future challenges". In: *IEEE Signal Processing Magazine* 31.1 (2013), pp. 34–44.

⁵ Nathan A Hagen and Michael W Kudenov. "Review of snapshot spectral imaging technologies". In: *Optical Engineering* 52.9 (2013), p. 090901.

⁶ Alex Pappachen James and Belur V Dasarathy. "Medical image fusion: A survey of the state of the art". In: *Information Fusion* 19 (2014), pp. 4–19.

by traditional spatio-spectral imaging sensors, resulting in high storage and processing costs.⁷

In other aspects, to reduce the high storage cost of the HS images, taking advantage of the Compressed sensing (CS) theory⁸, Compressive spectral imaging (CSI) has emerged as an acquisition framework that enables to sense and simultaneously reduce the data dimensionality without any further processing step.⁹ CSI allows retrieving the information from a small number of projections by assuming the sparsity property in the signal,¹⁰ in such a way that the cost of sensing, transmitting and processing of the HS image acquired using this approach is less than with traditional methods of acquisition.

In multi-resolution images fusion, once an image of high spatial and spectral resolution has been obtained, a common task to perform in this type of data is classification. Each spatial location in a spectral image is represented by a vector whose values correspond to the intensity of reflected light when illuminated with different wavelengths. These vectors are also known as the spectral signature of the pixels¹¹. In general, each material reflects a certain amount of electromagnetic energy when it is illuminated with specific wavelengths.

Therefore, the information provided by the spectral signatures makes it possible to distinguish different materials, which leads to the possibility of classify each pixel of the image in a group or class¹². This task can be categorized as supervised or unsupervised, depending on whether or not it involves a training stage. The training step, in the supervised method, consists of learning a quantity of labeled spectral signatures. Then, in the testing process, each pixel is assigned the label of the spectral signature that has less spectral difference¹³. However, there are applications in which the labeled samples are unavailable or difficult to acquire. For this reason, unsupervised techniques known as clustering are used, which groups a set of similar pixels without prior knowledge of the data.

Recent work has begun to address both approaches simultaneously: multi-resolution images fusion and CSI. In compressive spectral imaging the fusion has been addressed as

⁷ Naoto Yokoya, Claas Grohnfeldt, and Jocelyn Chanussot. "Hyperspectral and multispectral data fusion: A comparative review of the recent literature". In: *IEEE Geoscience and Remote Sensing Magazine* 5.2 (2017), pp. 29–56.

⁸ David L Donoho et al. "Compressed sensing". In: *IEEE Transactions on information theory* 52.4 (2006), pp. 1289–1306.

⁹ Gonzalo R Arce et al. "Compressive coded aperture spectral imaging: An introduction". In: *Signal Processing Magazine, IEEE* 31.1 (2014), pp. 105–115.

¹⁰ Henry Arguello and Gonzalo R Arce. "Colored coded aperture design by concentration of measure in compressive spectral imaging". In: *IEEE Transactions on Image Processing* 23.4 (2014), pp. 1896–1908.

¹¹ Gary A Shaw and Hsiao-hua K Burke. "Spectral imaging for remote sensing". In: *Lincoln Laboratory Journal* 14.1 (2003), pp. 3–28.

¹² Chein-I Chang. *Hyperspectral imaging: techniques for spectral detection and classification*. Vol. 1. Springer Science & Business Media, 2003.

¹³ James Dougherty, Ron Kohavi, and Mehran Sahami. "Supervised and unsupervised discretization of continuous features". In: *Machine Learning Proceedings 1995*. Elsevier, 1995, pp. 194–202.

CS with side information,¹⁴ where a similar signal or additional information is known and integrated into compressive reconstruction via convex minimization. Most approaches, use different concepts of side information to estimate the sparsity or its probability distribution to improve the quality recovery of the signal¹⁵. Specifically, in¹⁶, is proposed the development of an adaptive coded aperture design for spectral imaging with the compressive architecture single pixel camera, based on a multi-resolution approach. An RGB side image is used to define blocks of similar pixels, such that they can be used to design the coded aperture patterns.

Note that, the fusion works of multi-resolution compressive spectral images, reconstruct the merged data, but do not address later processing tasks. However, the reconstruction procedure is not strictly necessary to perform different processing tasks such as classification.

For this reason, taking into account that the multi-resolution images fusion is useful to improve the spatial resolution of the HS images, and additionally, that the compressive spectral technique allows reducing the dimensionality of the data, this work seeks to extract classification features by merging of information from an high-spatial-resolution image (RGB) with a high-spectral-resolution compressed measurements. The objective is to achieve a supervised classification of high accuracy.

Due to the CSI architectures often use 2D detectors,¹⁷ whose costs increase with the resolution, and that on other hand, the single pixel camera (SPC)¹⁸ has been recently employed as a low-cost CSI sensor¹⁹, in this work a single-pixel optical architecture is used as the sensor of high-spectral resolution measurements.

In detail, the development of this thesis is structured as follows: First, in section 1, general concepts of supervised classification of spectral images, the compressive spectral imaging method, and current multi-resolution image fusion techniques are introduced. Then, in section 2, a matrix model, some notation, a mathematical description, the flowchart, and several simulation results of a proposed supervised classification methodology to fuse the high spectral information of an HS image with the spatial information extracted

¹⁴ Edwin Vargas et al. "Spectral image fusion from compressive measurements". In: *IEEE Transactions on Image Processing* 28.5 (2019), pp. 2271–2282.

¹⁵ Jonathan Scarlett, Jamie S Evans, and Subhrakanti Dey. "Compressed sensing with prior information: Information-theoretic limits and practical decoders". In: *IEEE Transactions on Signal Processing* 61.2 (2012), pp. 427–439.

¹⁶ Hans Garcia et al. "Multi-Resolution Coded Apertures Based on Side Information for Single Pixel Spectral Reconstruction". In: *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE. 2018, pp. 2215–2219.

¹⁷ Xing Lin et al. "Dual-coded compressive hyperspectral imaging". In: *Optics letters* 39.7 (2014), pp. 2044–2047.

¹⁸ Marco F Duarte et al. "Single-pixel imaging via compressive sampling". In: *IEEE signal processing magazine* 25.2 (2008), pp. 83–91.

¹⁹ Hoover Fabián Rueda-Chacón, Cesar Augusto Vargas-García, and Henry Arguello-Fuentes. "Single-pixel optical sensing architecture for compressive hyperspectral imaging". In: *Revista Facultad de Ingeniería Universidad de Antioquia* 73 (2014), pp. 124–133.

from the superpixels of an RGB image is presented. After that, in section 3, are proposed a mathematical model to perform the supervised classification of spectral images acquired with single pixel optical architecture and side information from an RGB sensor, and a computational algorithm to solve the mathematical model. Additionally, the acquisition architecture implemented for the simultaneous capture of data with compressed measurements of an SPC and an RGB sensor is presented. In Section 4, to evaluate the performance of the proposed method, several simulation results from experimental and simulated data are presented. Finally, the conclusions of this work are summarized in section 5.

1. THEORETICAL FRAMEWORK

In this section, the theoretical framework related to this research work is presented, which addresses the spectral images, the task of supervised classification, the theory of compressive sensing, and the fusion of images of different resolution. In addition, the most recent advances in each of the topics exposed are described.

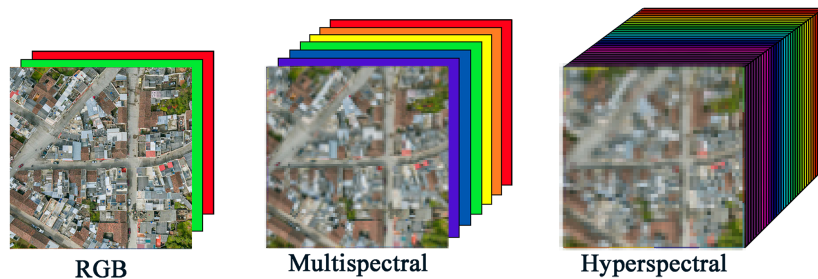
1.1. SUPERVISED CLASSIFICATION OF SPECTRAL IMAGING

1.1.1 Spectral Imaging. The electromagnetic spectrum is the range of electromagnetic radiation frequencies and their respective wavelengths. It is divided by frequencies into regions, bands or classes that receive different names: radio waves, microwaves, terahertz waves, infrared, visible light (380 to 950 nm), ultraviolet, X-rays, and gamma rays. The electromagnetic waves of each region have different characteristics, for example, how they are produced, how they interact with matter, and their practical applications²⁰.

The human visual system, in particular, detects and interprets the information from the visible light range²¹, the same spectral range that is captured by the traditional photo cameras. According to the colors that correspond to the wavelengths sensed by these optical devices (625, 525, and 470 nm), the acquired images are named RGB (red, green and blue).

In general, a multi-band image consists of a set of different spectral bands, covering a range of wavelengths. These images can be represented as a cube of three-dimensional data indexed by three exploratory variables (x , y , λ), where x and y are the two spatial dimensions of the scene, and λ is the spectral domain. Depending on the number of bands or spectral resolution, the multiband images can be of the RGB type composed of three spectral bands; multispectral (MS) formed by up to ten bands; or hyperspectral (HS) that contain up to hundreds of bands²². It is worth mentioning that HS images allow the acquisition of information of a greater spectral range (200 - 2500 nm) than the covered by RGB images. However, an important disadvantage of the devices that allow acquiring a more extensive spectral range is that their spatial resolution is usually limited. Fig. 1 shows an illustration of a multiband RGB, MS, and HS image.

Figure 1: Different types of multiband images according to their resolution



There are several methods to acquire a high-resolution spectral image (HS). For instance,

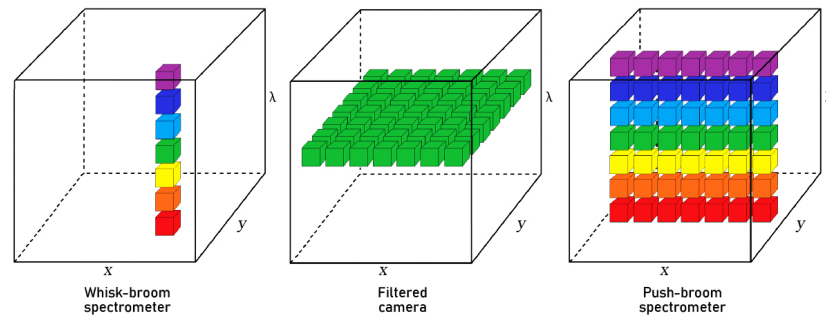
²⁰ Mark A Heald and Jerry B Marion. *Classical electromagnetic radiation*. Courier Corporation, 2012.

²¹ George Wald. "Human vision and the spectrum." In: *Science* (1945).

²² Chein-I Chang. *Hyperspectral data exploitation: theory and applications*. John Wiley & Sons, 2007.

an approach that captures the spectrum for a single point is called whiskbroom spectrometer; a filtered camera is other technique that acquires all the spatial pixels for each wavelength at the same time; and the push-broom spectrometer measures the spectrum through a spatial line and then scans the scene through the other spatial dimension⁵. Due to the scanning patterns of these conventional methods of spectral acquisition, based on the Nyquist-Shannon theory²³²⁴, the amount of information captured is abundant. For this reason, in addition to the low spatial resolution, the high storage cost is a limitation of the HS images. The traditional scanning methods mentioned above are illustrated in Fig. 2⁵.

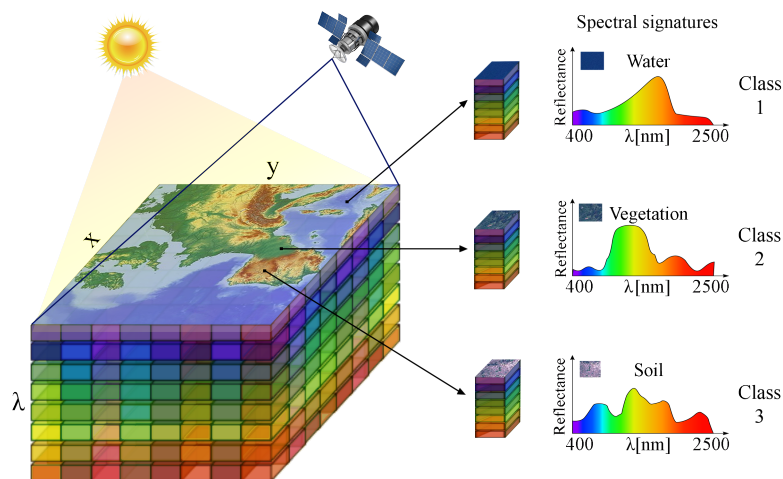
Figure 2: Traditional snapshot spectral imaging technologies



Source: *Review of snapshot spectral imaging technologies*, 2013⁵.

In a HS image, each spatial position is represented as a vector, known as a spectral signature, whose values correspond to the intensity at different spectral bands. In Fig. 3 the concept of spectral image and spectral signature is illustrated. Note that the curve formed by the amount of light reflected by a material (reflectance) when illuminated with different wavelengths is unique for each element.

Figure 3: Spectral imaging concept



²³ HJ Landau. "Sampling, data transmission, and the Nyquist rate". In: *Proceedings of the IEEE* 55.10 (1967), pp. 1701–1706.

²⁴ J Abdul. "The Shannon Sampling Theorem—Its Various Extensions and Applications: A". in: *Proceedings of the IEEE* 65.11 (1977), p. 1565.

Therefore, exploiting the fact that, all materials have unique spectral characteristics because they reflect, absorb, and emit radiation in a unique way, the information provided by the spectral signatures can be a valuable tool for tasks as the detection, identification, and classification of materials and objects with complex compositions³.

1.1.2 Supervised Classification. The classification of spectral images aims at assigning each spectral pixel to one class belonging to a set of categories based on its spectral characteristics. This task has found many applications in various fields such as military, precision agriculture, astronomy, and mineralogy and has led to a variety of methods²⁵. In the literature, many of the methods have concentrated on exploring the role of the spectral signatures in classification, employing exclusively the spectral signatures for determining the classification maps. Since spectral images can be seen as a set of 2D images acquired at different spectral wavelengths, the spatial information is naturally another source of data that can be considered to improve the classification performance. Then, the introduction of spatial dependency offers the possibility to boost the pixel-wise classification methods²⁶.

In general, there are two main approaches to the classification problem: supervised and unsupervised. Supervised techniques require the availability of a training set for learning the classifier. Then, in the testing process, each pixel is assigned the label of the spectral signature that has less spectral difference. However, there are applications in which the labeled samples are unavailable or difficult to acquire²⁷. For this reason, unsupervised techniques known as clustering are used, which groups a set of similar pixels without knowledge of the data. In general, in those applications where it is possible to have a set of sample training labels, it is preferred to use a supervised classification method. Multiple works show that this approach is what allows to achieve more accurate classification results.

There are many algorithms to perform the task of supervised classification: support vector machine (SVM)²⁸, k-nearest neighbor (KNN)²⁹, maximum likelihood, spectral angle mapping (SAM), artificial neural network (ANN), and decision tree, among others.

In particular, SVMs have shown a good performance for spectral image classification³⁰. In this, a few spectral signatures are used as training samples to train the classifier and then, the remaining spectral signatures are classified. The SVM algorithm aims to discriminate two classes by fitting an optimal separating hyperplane to the training data within a multi-dimensional feature space, by using only the closest training samples. Thus, the approach only considers samples close to the class boundary and work well

²⁵ Pedram Ghamisi et al. "Advanced spectral classifiers for hyperspectral images: A review". In: *IEEE Geoscience and Remote Sensing Magazine* 5.1 (2017), pp. 8–32.

²⁶ Mathieu Fauvel et al. "Advances in spectral-spatial classification of hyperspectral images". In: *Proceedings of the IEEE* 101.3 (2012), pp. 652–675.

²⁷ Claire A Baldeck and Gregory P Asner. "Estimating vegetation beta diversity from airborne imaging spectroscopy and unsupervised clustering". In: *Remote Sensing* 5.5 (2013), pp. 2057–2071.

²⁸ J Anthony Gualtieri and Robert F Crompt. "Support vector machines for hyperspectral remote sensing classification". In: *The 27th AIPR Workshop: Advances in Computer-Assisted Recognition*. International Society for Optics and Photonics. 1999, pp. 221–232.

²⁹ James M Keller, Michael R Gray, and James A Givens. "A fuzzy k-nearest neighbor algorithm". In: *IEEE transactions on systems, man, and cybernetics* 4 (1985), pp. 580–585.

³⁰ Phan Thanh Noi and Martin Kappas. "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery". In: *Sensors* 18.1 (2018), p. 18.

with small training sets, even when high dimensional data sets are classified³¹. Variations of the SVM-based algorithms have also been proposed to improve the classification accuracy. These variations include semisupervised learning which exploits both labeled and unlabeled samples³², postprocessing of the individually labeled samples based on certain decision rules³³, and incorporating spatial information directly in the SVM kernels³⁴.

On the other hand, the KNN classifier, in addition to taking into account the similarity between the pixel and the training signatures, finds which is the class that occurs most frequently in the neighborhood to assign the label to the pixel. The SVM and KNN are two of the most well-known, studied and tested algorithms in the area of spectral imaging. A detailed mathematical description of these two classifiers is developed in section 2.2.

1.2. COMPRESSIVE SPECTRAL IMAGING

Most of the information acquired by traditional sampling methods of spectral images is discarded during the compression process before being stored. Due to the high dimensionality of the spectral images, recent works have focused on obtaining a compressed signal or spectral image directly avoiding the compression step.

The theory of compressive sampling (CS) establishes that it is possible to recover images or signals from fewer samples than those required by the Nyquist/Shannon criterion⁵. Specifically, CSI captures 2D coded and dispersed projections of the 3D scene rather than direct measurements of the voxels. Therefore, a far less number of samples are acquired compared to traditional techniques, leading to an improved sensing speed⁹. CSI theory establishes that an estimation of the 3D data cube can be successfully recovered from the compressed measurements. To make this possible, CSI relies on two principles: sparsity, which characterizes the spectral scene of interest, and incoherence, which shapes the sensing structure⁸.

A signal is K -sparse if at most K of its components are nonzero. However, many natural and man-made signals are not sparse but compressible in the sense that they can be well-approximated as a linear combination of just a few elements from a known basis or dictionary Ψ . Formally, denote $\mathcal{F}$ as the spatio-spectral input data cube, with $M \times N$ spatial dimensions, and L spectral bands. Then, a spectral image can be expanded in an orthonormal basis $\Psi = [\psi_1, \psi_2, \dots, \psi_{MNL}]$ as

$$\bar{\mathbf{f}} = \sum_{i=1}^{MNL} \theta_i \psi_i, \quad (1.1)$$

where $\bar{\mathbf{f}} \in \mathbb{R}^{MNL}$ is a vector representation of the source $\mathcal{F}$, θ is the coefficient sequence

-
- ³¹ Mathieu Fauvel et al. "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles". In: *IEEE Transactions on Geoscience and Remote Sensing* 46.11 (2008), pp. 3804–3814.
 - ³² Lorenzo Bruzzone, Mingmin Chi, and Mattia Marconcini. "A novel transductive SVM for semisupervised classification of remote-sensing images". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.11 (2006), pp. 3363–3373.
 - ³³ Francesca Bovolo et al. *A novel context-sensitive SVM for classification of remote sensing images*. University of Trento, 2006.
 - ³⁴ Gustavo Camps-Valls et al. "Composite kernels for hyperspectral image classification". In: *IEEE geoscience and remote sensing letters* 3.1 (2006), pp. 93–97.

of $\mathbf{f}$, such that $\theta_i = \langle \bar{\mathbf{f}}, \psi_i \rangle$. Then, the spectral image $\bar{\mathbf{f}}$ can be expressed as $\bar{\mathbf{f}} = \Psi\theta$, where Ψ is an $MNL \times MNL$ matrix with $\psi_1, \psi_2, \dots, \psi_{MNL}$ as columns.

On the other hand, let $\mathbf{H} \in \mathbb{R}^{MV \times MNL}$, where MV is the number of compressed measurements to acquire, and $\Psi \in \mathbb{R}^{MNL \times MNL}$ be two orthonormal bases, the CSI incoherence property measures how correlated are the elements of $\mathbf{H}$ and Ψ where a low correlation is desired³⁵. Following the compressive sensing theory, the first basis is used to measure/sense the signal $\bar{\mathbf{f}}$, and the second basis is used to represent $\bar{\mathbf{f}}$ in a sparse domain. The mutual coherence of the orthonormal basis $\mathbf{H}$ and Ψ is defined as the maximum absolute value of the inner product between any two columns of the basis, given by

$$\mu(\mathbf{H}, \Psi) = \sqrt{MN} \max_{1 \leq i, j \leq MN} |\langle \mathbf{H}_i, \Psi_j \rangle|, \quad (1.2)$$

where $\mathbf{H}_i$ represents the i -th column of the sensing matrix $\mathbf{H}$, whose structure is determined by the sensing architecture. The coherence measures the maximum correlation between any two elements of $\mathbf{H}$ and Ψ . In CSI a minimum coherence among the orthonormal base is desired.

Coded Apertures or Coded-Aperture Masks are grids, gratings, or other patterns of materials opaque to various wavelengths of light. A coded aperture can be represented as a binary matrix, where the one-valued elements allow the electromagnetic radiation to pass through and the zero-valued elements block the light, generating a coding pattern in the object of interest.

In the CSI mathematical model, the spectral image $\mathcal{F}$ with $M \times N$ spatial dimensions and L spectral bands is discretized as $\mathcal{F}_{m,n,k}$, where m and n index the spatial coordinates, and k determines the k -th spectral band. $\mathcal{F}$ can also be represented in vector form as $\mathbf{f} = [\mathbf{f}_0^T, \dots, \mathbf{f}_{L-1}^T]^T$, where each spectral band $\mathbf{f}_k$ can be expressed as $\mathbf{f}_k = [\mathcal{F}_{0,0,k}, \mathcal{F}_{1,0,k}, \dots, \mathcal{F}_{(M-1),0,k}, \dots, \mathcal{F}_{0,1,k}, \mathcal{F}_{1,1,k}, \dots, \mathcal{F}_{(M-1),1,k}, \dots, \mathcal{F}_{(M-1),(N-1),k}]^T$.

In general, CSI projections can be written in matrix notation as

$$\mathbf{y} = \mathbf{H}\mathbf{f} = \mathbf{H}\Psi\theta, \quad (1.3)$$

where $\mathbf{y}$ is the measurement set in vector form and, $\mathbf{H}$ is known as the system sensing matrix whose entries are determined by the CSI optical setup⁹. The spectral image $\mathbf{f}$ is said to have a sparse representation θ in a basis Ψ .

Although CSI allows to reconstruct a spectral image from a considerable smaller number of samples than those required by the Shannon-Nyquist theorem, there is a minimum number of necessary measurements that must be acquired to successfully reconstruct a scene. Considering a spectral image of size MNL which is K -sparse in a basis Ψ , the CSI theory states that $u \geq K \log(MNL)$ measurements are necessary in order to achieve a correct reconstruction. However, for spectrally rich scenes or very detailed spatial scenes, a single shot CSI measurement may not provide a sufficient number of compressed measurements to satisfy the above constraint. This leads to an excessively undetermined inverse problem hence failing to obtain an adequate reconstruction. Increasing the number of measurement shots will multiply the number of measurements,

³⁵ David L Donoho and Xiaoming Huo. "Uncertainty principles and ideal atomic decomposition". In: *Information Theory, IEEE Transactions on* 47.7 (2001), pp. 2845–2862.

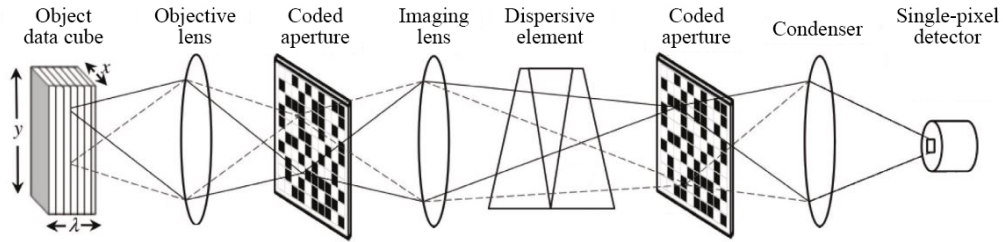
thus rapidly overcoming such limitations ³⁶.

The inverse problem $\mathbf{f} = \mathbf{H}^{-1}\mathbf{y}$ would be ill-conditioned, because it has infinite solutions, due to the smaller number of measurements compared to the signal dimension. For this reason, this problem can be solved via optimization algorithms that obtain a θ approximation from $\hat{\mathbf{f}} = \Psi\{\arg \min_{\theta} \|\mathbf{H}\Psi\theta - \mathbf{y}\|_2^2 + \tau\|\theta\|_1^2\}$, where τ is a regularization constant.

In recent decades, optical architectures have been developed to implement CS for spectral image acquisition. For example, the Spatial-spectral encoded compressive HS imaging (SSCSI) ³⁷, the Coded aperture snapshot spectral imaging (CASSI) ⁹, the Dual-coded hyper-spectral imaging (DCSI) ¹⁷, the Prism-mask multispectral video imaging system (PMVIS) ³⁸ and the aforementioned Single-pixel camera (SPC), which have shown to substantially improve the performance, the data acquisition times and the processing capacity when configured via compressive sampling ¹⁸.

In particular, the SPC employs a single pixel detector to capture the spatial/spectral scene. Different variations have already been proposed, including the single pixel compressive hyper spectral imaging sensor (SPHIS), which consists of two coded apertures to spatially encode the scene, a prism or dispersive element, an energy condenser and a single-pixel detector ¹⁹ as is depicted in Fig. 4. Another configuration of the Single-

Figure 4: Optical architecture of the SPHIS.



Source: *Single-pixel optical sensing architecture for compressive hyperspectral imaging*, 2014 ¹⁹.

pixel camera uses a spectrometer, instead of a single pixel detector, which reduces the optical complexity of the system. By replacing the single pixel sensor with a spectrometer the two-dimensional spatial plane can be extended to the data-cube (Fig. 5). All SPC variations use a micro-mirror (digital micro-mirror device-DMD) array for spatial light modulation. The modulation is driven by a pseudo-random spatial code and a single optical measurement is taken per each modulation. The system forms an image on the DMD, then the reflected modulated light from the DMD is focused on the spectrometer and measured. The spectrometer captures the spectral information in different bins, representing the spectral bands.

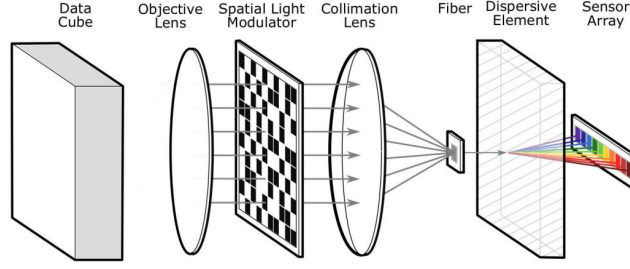
In classification, a spectral signature is assumed to be a linear combination of very few atoms from an overcomplete dictionary consisting of a set of training data from all the

³⁶ Carlos A Hinojosa et al. *Compressive spectral imaging using multiple snapshot colored-mosaic detector measurements*. 2016. DOI: [10.1117/12.2224369](https://doi.org/10.1117/12.2224369).

³⁷ Xing Lin et al. "Spatial-spectral encoded compressive hyperspectral imaging". In: *ACM Transactions on Graphics (TOG)* 33.6 (2014), p. 233.

³⁸ Xun Cao et al. "A prism-mask system for multispectral video acquisition". In: *IEEE transactions on pattern analysis and machine intelligence* 33.12 (2011), pp. 2423–2435.

Figure 5: Single pixel combined with spectrometer.



classes. The spectral signatures of pixels belonging to the same class are assumed to approximately lie in a low-dimensional subspace. Suppose we have C distinct classes and the c -th class has N_c training samples $\mathbf{d}_{n_c}^c = \{\mathbf{d}_{n_c}^c \in \mathbb{R}^{K \times 1} | 1 \leq n_c \leq N_c\}$. If $\mathbf{x}_{i,j}$ belongs to the c -th class, then its spectrum approximately lies in a low-dimensional subspace spanned by the training samples in the c -th class. Based on the above sparsity assumption, an unknown test sample is modeled to lie in the union of the C subspaces associated with the C classes. Let $\mathbf{D}^c = [\mathbf{d}_1^c \dots \mathbf{d}_{N_c}^c]$ with $\mathbf{D}^c \in \mathbb{R}^{K \times N_c}$ be a dictionary and by combining the class subdictionaries, a test sample $\mathbf{x}_{i,j}$ can be written as a sparse linear combination of all the training samples as

$$\mathbf{x}_{i,j} = \sum_{c=1}^C \mathbf{D}^c \mathbf{s}_{i,j}^c, \quad (1.4)$$

where $\mathbf{s}_{i,j}^c \in \mathbb{R}^{N_c \times 1}$ is a sparse vector associated with each class subdictionary. Ideally, if $\mathbf{x}_{i,j}$ belongs to the c -th class, then $\mathbf{s}_{i,j}^{\bar{c}} = \mathbf{0}$, where $1 \leq \bar{c} \leq C$ and $\bar{c} \neq c$.

In ³⁹, compressive supervised spectral image classification is formulated, based on finding the sparse representation of each pixel in a training sample basis directly from a set of compressive measurements. Then, the Equation 1.4 for all pixels can now be expressed as

$$\begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_{IJ}^T \end{bmatrix} = \begin{bmatrix} (\mathbf{s}_1^1)^T & \dots & (\mathbf{s}_1^C)^T \\ \vdots & \ddots & \vdots \\ (\mathbf{s}_{IJ}^1)^T & \dots & (\mathbf{s}_{IJ}^C)^T \end{bmatrix} \begin{bmatrix} (\mathbf{D}^1)^T \\ \vdots \\ (\mathbf{D}^C)^T \end{bmatrix}, \quad (1.5)$$

where $\mathbf{s}_n^c \in \mathbb{R}^{N_c \times 1}$ is the representation of the c -th class in the n -th pixel. By using a tensor representation, the classification model is now given by

$$\mathbf{x} = ([\mathbf{D}^1, \dots, \mathbf{D}^C] \otimes \mathbf{I}_{IJ}) \text{vec} \left(\begin{bmatrix} (\mathbf{s}_1^1)^T & \dots & (\mathbf{s}_1^C)^T \\ \vdots & \ddots & \vdots \\ (\mathbf{s}_{IJ}^1)^T & \dots & (\mathbf{s}_{IJ}^C)^T \end{bmatrix} \right) = \bar{\mathbf{D}} \mathbf{s}, \quad (1.6)$$

where $\mathbf{I}_{IJ}$ is the identity matrix with dimensions $IJ \times IJ$. Based on the advantages of the structure of the $\bar{\Phi}$ matrix as the measurement basis, the following representation for

³⁹ Ana Ramirez et al. "Spectral image classification from optimal coded-aperture compressive measurements". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6 (2014), pp. 3299–3309.

compressive classification based CS is formulated

$$\mathbf{y} = \overline{\Phi}\mathbf{x} + \omega = \overline{\Phi}\overline{\mathbf{D}}\mathbf{s} + \omega. \quad (1.7)$$

Then, given a structured dictionary $\overline{\mathbf{D}}$, the unknown sparse vector $\mathbf{s}$ satisfying $\mathbf{y} = \overline{\Phi}\overline{\mathbf{D}}\mathbf{s}$ can be obtained from the following sparsity optimization:

$$\min_{\mathbf{s}} \quad (1/2)\|\overline{\Phi}\overline{\mathbf{D}}\mathbf{s} - \mathbf{y}\|_2^2 + \lambda \sum_{c=1}^C \|\mathbf{s}^c\|_1. \quad (1.8)$$

Subsequently, given the estimated sparse vector for all the pixels $\mathbf{s} = [\mathbf{s}^1, \dots, \mathbf{s}^C]$, the classification method labels each pixel in one of the known classes. Assuming that there are C different classes and that every single observation pixel belongs to one of the given classes, the problem of classifying each test pixel consists of finding the class whose training pixel is the nearest to the test pixel in the Euclidean distance sense.

1.3. MULTI-RESOLUTION IMAGE FUSION

As a consequence of the low spatial resolution of this type of data, the problem of merging a high spectral and low spatial resolution image with an auxiliary image of higher spatial resolution but lower spectral resolution, also known as multi-resolution image fusion, has been explored for many years and is still an active, challenging and innovative field of research ⁴⁰.

Data fusion techniques provide a methodological framework for finding a high spatial and spectral resolution image. Multiple algorithms have been developed to solve the data fusion problem, which based on the used theoretical approach can be categorized as component substitution, spectral unmixing, sparse signal representation, and Bayesian estimation ⁴¹.

In general, data fusion techniques are typically evaluated on images captured by traditional spectral imaging sensors, whose acquisition is based on the Nyquist-Shannon sampling theorem. More precisely, these sensing systems acquire the information of the scene of interest using a constant sampling rate along the three dimensions, leading to large amounts of information, high acquisition times and expensive processing cost ⁴².

When considering remotely sensed images, an archetypal fusion task is the *pansharpening* which generally consists of fusing a high spatial resolution panchromatic (PAN) image and low spatial resolution MS image. Pansharpening has been addressed in the literature for several decades and still remains an active topic. An example is shown in Fig. 6.

In ⁴², the fusion is addressed from Bayesian methods. In this work are studied new multi-band image fusion algorithms to enhance the resolution of HS image. For instance, a hierarchical Bayesian framework is proposed for multi-band image fusion by incorporating forward model, statistical assumptions and Gaussian prior for the target image to be restored. In addition, to derive Bayesian estimators associated with the resulting posterior distribution, two algorithms based on Monte Carlo sampling and optimization strategy are also developed. Indeed, solving a Sylvester equation and projecting onto a simplex resulting from the non-negativity and sum-to-one constraints other fusion algorithm is proposed.

Recent works has begun to address the multi-resolution images fusion acquired with some compressive spectral architectures. In compressive spectral imaging the fusion has been addressed as CS with side information ⁴³, where a similar signal or additional information is known and integrated into compressive reconstruction via convex minimization. Most approaches, use different concepts of side information to estimate the sparsity or its probability distribution to improve the quality recovery of the signal ¹⁵. Specifically, in the work of ¹⁶, is proposed the development of an adaptive coded aperture design for spectral imaging with the compressive architecture single pixel camera based on a multi-

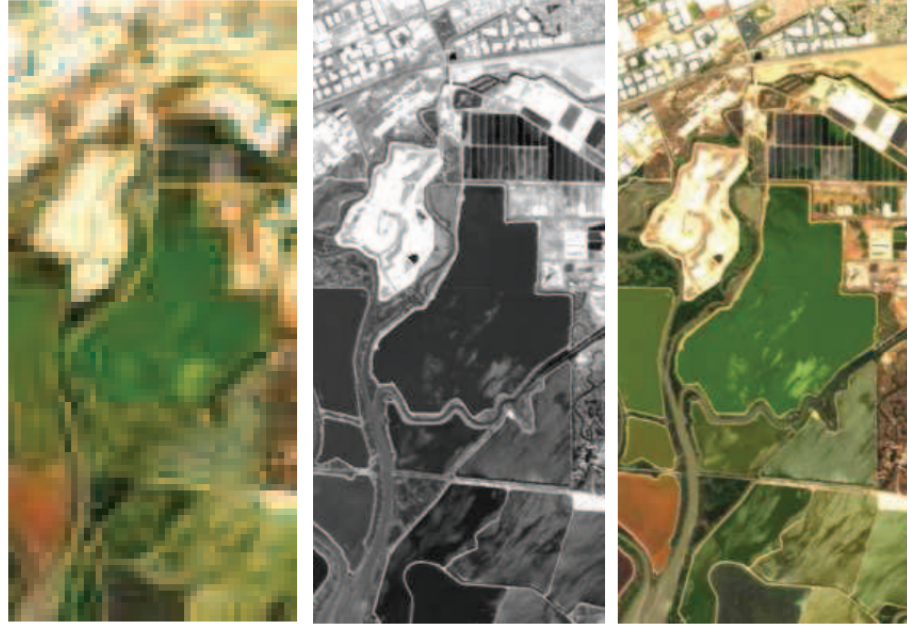
⁴⁰ Laetitia Loncan et al. "Hyperspectral pansharpening: A review". In: *arXiv preprint arXiv:1504.04531* (2015).

⁴¹ Q. Wei et al. "Multiband Image Fusion Based on Spectral Unmixing". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.12 (2016), pp. 7236–7249. DOI: [10.1109/TGRS.2016.2598784](https://doi.org/10.1109/TGRS.2016.2598784).

⁴² Qi Wei. "Bayesian fusion of multi-band images: a powerful tool for super-resolution". PhD thesis. Institut national polytechnique de Toulouse (INPT), 2015.

⁴³ Xin Yuan, Yangyang Sun, and Shuo Pang. "Compressive video sensing with side information". In: *Applied optics* 56.10 (2017), pp. 2697–2704.

Figure 6: (Left) Hyperspectral image (size: $99 \times 46 \times 224$, res.: $80\text{m} \times 80\text{m}$). (Middle) Panchromatic image (size: 396×184 res.: $20\text{m} \times 20\text{m}$). (Right) Target image (size: $396 \times 184 \times 224$ res.: $20\text{m} \times 20\text{m}$).



Source: *Bayesian fusion of multi-band images: a powerful tool for super-resolution*, 2014 ⁴².

resolution approach. An RGB side image is used to define blocks of similar pixels, such that they can be used to design the coded aperture patterns.

In ¹⁴, the compressive fusion process is formulated as an inverse problem that minimizes an objective function defined as the sum of a quadratic data fidelity term and smoothness and sparsity regularization penalties. The parameters of the different sensors are optimized and the choice of an appropriate regularization is studied in order to improve the quality of the high resolution reconstructed images.

Note that, the fusion works of multi-resolution compressive spectral images, reconstruct the merged data, but do not address later processing tasks. However, the reconstruction procedure is not strictly necessary to perform different processing tasks such as classification.

2. SUPERVISED SPATIO-SPECTRAL CLASSIFICATION OF FUSED IMAGES USING SUPERPIXELS

In the next pages, is presented a supervised classification approach that merges the spatial information of superpixels extracted from a RGB image with the spectral information of an HS image. This fusion procedure is formulated as an optimization problem, which is solved using an alternating direction method of multipliers (ADMM), and then a supervised classifier is used. In general, the framework developed provides a faster way to directly fuse and classify hyperspectral images, with better overall accuracy compared with other fusion methods. This approach reduces the classification time since the image is grouped in superpixels, then the number of pixels that need to be classified is less.

The mathematical model shown in this section was presented at the OSA Imaging and Applied Optics Congress in 2018, USA, paper “*Supervised Classification of Hyperspectral Images using Side Information*”⁴⁴. In addition, the full version presented here was published in the Volume 58, Issue 7, pp. B9-B18 of the Applied Optics journal by The Optical Society of America, USA, 2019, Category A1, paper “*Supervised spatio-spectral classification of fused images using superpixels*”⁴⁵.

2.1. MATRIX MODEL

In general, denote $\mathcal{F}$ as the spatio-spectral datacube, with $M \times N$ spatial dimensions, L spectral bands, and entries denoted as $\mathcal{F}_{n_1, n_2, k}$, where n_1 and n_2 index the spatial coordinates, and k determines the k -th spectral band. The datacube can be rearranged as a matrix $\mathbf{F} \in \mathbb{R}^{MN \times L}$ whose rows $\{\mathbf{f}_j^T\}_{j=0}^{MN}$ are the spectral signatures of the datacube. Specifically, the rearrangement can be expressed as

$$\mathbf{F} = \begin{bmatrix} [\mathcal{F}_{0,0,0}, \mathcal{F}_{1,0,0}, \dots, \mathcal{F}_{0,1,0}, \dots, \mathcal{F}_{(M-1),(N-1),0}]^T, \\ \dots, [\mathcal{F}_{0,0,(L-1)}, \dots, \mathcal{F}_{(M-1),(N-1),(L-1)}]^T \end{bmatrix}. \quad (2.1)$$

In this work, the classification maps of a high spatial and spectral resolution image $\mathbf{F}$ are found from fused image obtained from an RGB and an HS image, denoted as $\mathbf{F}_m \in \mathbb{R}^{MN \times 3}$ and $\mathbf{F}_h \in \mathbb{R}^{M_h N_h \times L}$, respectively. Note that $\mathbf{F}_m$ can be composed by combining three spectral bands representing the red, green, and blue colors. Further, $\mathbf{F}_h$ can be seen as a blurred and spatially downsampled instance of $\mathbf{F}$ with $N_h = N/p$ and $M_h = M/p$, where p is the downsampling factor along the spatial coordinates. Mathematically, the matrices $\mathbf{F}_m$ and $\mathbf{F}_h$ can be expressed as

$$\mathbf{F}_m = \mathbf{F} \mathbf{D}_m, \quad (2.2)$$

$$\mathbf{F}_h = \mathbf{D}_h \mathbf{F}, \quad (2.3)$$

⁴⁴ Karen Sanchez, Carlos Hinojosa, and Henry Arguello. “Supervised Classification of Hyperspectral Images using Side Information”. In: *Laser Applications to Chemical, Security and Environmental Analysis*. Optical Society of America. 2018, JW5E–5.

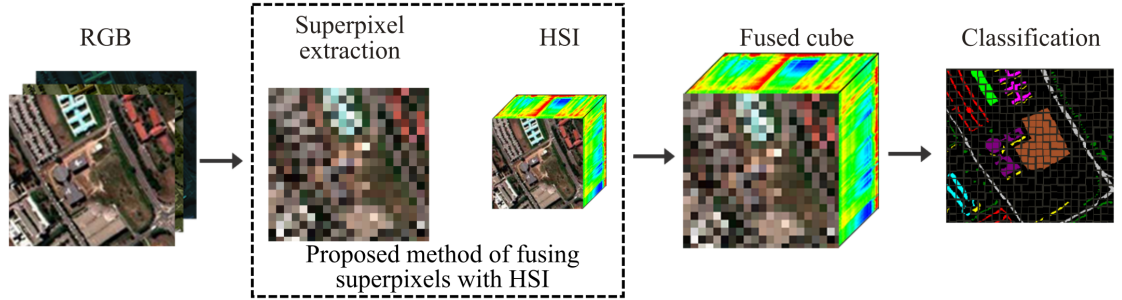
⁴⁵ Karen Sanchez, Carlos Hinojosa, and Henry Arguello. “Supervised spatio-spectral classification of fused images using superpixels”. In: *Applied optics* 58.7 (2019), B9–B18.

where $\mathbf{D}_m \in \mathbb{R}^{L \times 3}$ is a sparse selection matrix with only three nonzero values, i.e., a single one per column that indicates the selected spectral bands composing the RGB image and $\mathbf{D}_h \in \mathbb{R}^{M_h N_h \times MN}$ is a spatial downsampling that includes both the blur operation and the spatial decimation of $\mathbf{F}$ with downsampling factor p .

2.2. PROPOSED SPATIO-SPECTRAL CLASSIFICATION APPROACH

The aim of the proposed method is to incorporate the spatial information of the spectral scene in the classification features using an RGB image as side information. The flowchart of the algorithm is depicted in Fig. 7. In general, a rich spatial and spectral fused cube $\mathbf{F} \in \mathbb{R}^{M_m N_m \times L_h}$ is obtained using the spectral information of a HS image and the superpixels extracted from the RGB image. Then, the fused cube is later used as the input for a supervised pixel-based classifier.

Figure 7: Flowchart of the proposed classification framework. An RGB image is segmented into a predefined number of superpixels. The superpixel image and the HS image are used to obtain a high spectral and spatial resolution which is used as classification features. Finally, a supervised classification algorithm is applied on the fused image.



2.2.1. Superpixel Segmentation. In Fig. 7, the first step in the proposed methodology is to segment an RGB image in superpixels. The superpixels technique is based on the over-segmentation strategy that gather a group of uniform pixels into homogeneous subregions. In this work, we use the simple linear iterative clustering (SLIC) superpixel algorithm Xuewen Zhang et al. “SLIC superpixels for efficient graph-based dimensionality reduction of hyperspectral imagery”. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*. vol. 9472. International Society for Optics and Photonics. 2015, p. 947209; Radhakrishna Achanta et al. “SLIC superpixels compared to state-of-the-art superpixel methods”. In: *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012), pp. 2274–2282. SLIC works in the 5-D space, where two coordinate components (x, y) depict the spatial location of the segment and the other three components corresponds to the RGB color channels. Given N_{seg} desired equally-sized superpixels, where the approximation size of each superpixel is N^2/N_{seg} , the first step of SLIC is to define a cluster center at every grid interval $S = \sqrt{N^2/N_{seg}} = N/\sqrt{N_{seg}}$. The algorithm assumes that the pixels associated with a cluster, lie in a $2S \times 2S$ area around the superpixel center on the (x, y) plane. Therefore, this becomes the search area for the pixels near to each cluster center. To avoid keeping the center on the edge of an object, it is transferred to the lowest gradient position in a 3×3 neighborhood. In the next step, for each cluster center, SLIC assigns the best matching pixels from the search area according to the distance measure

$$D_c = \sqrt{(R_j - R_{j'})^2 + (G_j - G_{j'})^2 + (B_j - B_{j'})^2}, \quad (2.4)$$

$$D_p = \sqrt{(x_j - x_{j'})^2 + (y_j - y_{j'})^2}, \quad (2.5)$$

$$D_t = D_c + \frac{m}{S} D_p, \quad (2.6)$$

where R_j, G_j, B_j corresponds to the color of the j -th pixel, and m controls the compactness of a superpixel, which is usually chosen as $m = 10$. The complexity of SLIC algorithm is $O(MN)$, since it needs to compute distances from any point to no more than eight cluster centers and the number of iterations is constant⁴⁶. Notice that the region grows at most two times the cluster radius with the SLIC algorithm.

Once the segmentation map of the RGB image has been created using SLIC, the next step in the proposed method is to create the upsampling matrix $\mathbf{U} \in \mathbb{R}^{N_{seg} \times MN}$ which collects the superpixels information. First, denote $\mathbf{p}^e$ as the vector of size n_e containing the indexes of all pixels belonging to the superpixel e . Then, the nonzero values of the e -th row of $\mathbf{U}$, denoted in vector form as $(\mathbf{u}_e)^T$, are determined by the entries of $\mathbf{p}^e$ and the value of n_e as follows

$$(\mathbf{u}_e)_{(\mathbf{p}^e)_l}^T = \frac{1}{n_e}, \quad \text{for } l = 0, \dots, n_e, \quad (2.7)$$

where $(\mathbf{u}_e)_{(\mathbf{p}^e)_l}^T$ denotes the position in $\mathbf{u}_e$ indexed by the l -th entry of the $\mathbf{p}^e$ vector.

2.2.2. Problem Formulation. The following step in the proposed method is to perform the fusion operation $\mathbf{F}_h$ spectral information and the superpixels extracted from the RGB image $\mathbf{F}_m$. Using the notation given in section 2.1, the proposed fusion operation is formulated as the optimization problem

$$\min_{\mathbf{F}} \frac{1}{2} \|\mathbf{F}_h - \mathbf{D}_h \mathbf{F}\|_F^2 + \lambda \|\mathbf{F}\|_*, \quad (2.8)$$

where $\|\cdot\|_F$ stands for the Frobenius norm and λ is a regularization parameter. Since pixels in $\mathbf{F}$ that belongs to the same class share a common low-rank pattern, the nuclear norm minimization $\|\mathbf{F}\|_*$, in Eq. (2.8), is used to obtain the lowest-rank representation that captures the global structure of the data⁴⁶.

As the matrix $\mathbf{F}$ is obtained from the superpixels information, it is possible to relate the $\mathbf{U}$ matrix with $\mathbf{F}$ as

$$\mathbf{F} = \mathbf{U}^T \bar{\mathbf{F}}, \quad (2.9)$$

where $\bar{\mathbf{F}} \in \mathbb{R}^{N_{seg} \times L}$ contains the fused spectral information of each superpixel. Then, assuming that $\text{rank}(\bar{\mathbf{F}}) \approx \text{rank}(\mathbf{F})$, the optimization problem in Eq. (2.8) can be rewritten as

$$\min_{\bar{\mathbf{F}}} \frac{1}{2} \|\mathbf{F}_h - \mathbf{D}_h \mathbf{U}^T \bar{\mathbf{F}}\|_F^2 + \lambda \|\bar{\mathbf{F}}\|_*. \quad (2.10)$$

A solution for this optimization problem is shown in Appendix A.

2.2.3. Supervised Classification. Given the spectral superpixels $\bar{\mathbf{F}}$ and the respective high spectral and spatial resolution image $\mathbf{F}$, obtained with Eq. (2.9), the multiclass supervised classification method aims at determining the class label that better fits the

⁴⁶ Zhi He et al. "Low-rank group inspired dictionary learning for hyperspectral image classification". In: *Signal Processing* 120 (2016), pp. 209–221.

corresponding j_s -th pixel, for $j_s = 0, \dots, N_{seg} - 1$. Notice that with the proposed method it is only necessary to determine the class label of the N_{seg} spectral superpixels (rows of $\bar{\mathbf{F}}$). Then, the pixels in $\mathbf{F}$ acquire the same label of the superpixel which they belong to. As the number of superpixels $N_{seg} \ll MN$, the proposed approach boosts the computational performance, as the next section shows. In this paper, the classification of the spectral scene is achieved using two well-known algorithms: the support vector machine (SVM) and the k -nearest neighbor (KNN). In addition, we denote $\Omega = \{\omega_1, \dots, \omega_c\}$ as the set of c class labels and $\Theta = \{\theta_1, \dots, \theta_v\}$ as the set of v integers indexing the training samples.

2.2.3.1. SVM. Let first consider the binary supervised classification problem, where each feature vector of the training set is associated with a binary output $\beta_\theta \in \{-1, 1\}$ for $\Theta = \{\theta_1, \dots, \theta_v\}$. In this context, the SVM focuses on determining a hyperplane that optimally separates the samples belonging to different classes. This problem is equivalent to minimizing the norm of a weight vector $\mathbf{h}$ that is normal to the decision hyperplane^{47,48}. Specifically, this classification method solves the following optimization problem

$$\begin{aligned} \min_{\mathbf{h}, h_0, \mathcal{X}_\Theta} \quad & \left\{ \frac{1}{2} \|\mathbf{h}\|_2^2 + \eta \sum_{\Theta} \mathcal{X}_\Theta \right\} \\ \text{s.t.} \quad & \beta_\Theta \left(\gamma(\mathbf{f}_\Theta^T)^T \mathbf{h} + h_0 \right) \geq 1 - \mathcal{X}_\Theta \\ & \mathcal{X}_\Theta \geq 0, \end{aligned} \quad (2.11)$$

for all $\Theta = \{\theta_1, \dots, \theta_v\}$, where $\mathbf{f}_\Theta^T$ are the Θ rows of $\mathbf{F}$ used as training data; $\mathcal{X}_\Theta$ is the set of slack variables that considers the non-separability between sets belonging to different classes; η is the regularization parameter that controls the influence of the misclassified samples; γ is a non-linear function that maps the feature vector to the kernel space; h_0 is the offset of the decision hyperplane with respect to the origin of the coordinate system⁴⁹.

As mentioned before, SVM are basically binary classifiers. However, a multiclass strategy for the spectral image classification is required. To overcome this drawback, the one-against-one multiclass strategy is adopted, where two classes are separately analyzed by implementing a binary classifier, ignoring, at the same time, the remaining classes⁵⁰. Finally, using the training sample set $\{\mathbf{f}_\Theta^T\}_{\Theta=\theta_1}^{\theta_v}$, the problem of classifying each test pixel consists of finding the class Ω whose training pixel Θ is the nearest to the test pixel in the Euclidean distance

$$\text{Class}(\bar{\mathbf{f}}_{j_s}^T) = \min_{\Omega=\Omega_1, \dots, \Omega_c} \|\bar{\mathbf{f}}_{j_s}^T - (\mathbf{f}_\Theta^T)_{(\Omega)}\|_2^2, \quad (2.12)$$

where $\bar{\mathbf{f}}_{j_s}^T$ is the j_s -th row of $\bar{\mathbf{F}}$ and $(\mathbf{f}_\Theta^T)_{(\Omega)}$ denotes the training pixels (rows of $\mathbf{F}$) belonging to the class Ω . Notice that the SVM training is performed using the rows of $\mathbf{F}$ as training

⁴⁷ Giorgos Mountrakis, Jungho Im, and Caesar Ogoe. "Support vector machines in remote sensing: A review". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 66.3 (2011), pp. 247–259.

⁴⁸ Jun Li et al. "Generalized composite kernel framework for hyperspectral image classification". In: *IEEE transactions on geoscience and remote sensing* 51.9 (2013), pp. 4816–4829.

⁴⁹ Bernhard Schölkopf, Alexander J Smola, and Francis Bach. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. The MIT Press, 2018.

⁵⁰ Bin Gu and Victor S Sheng. "A Robust Regularization Path Algorithm for ν -Support Vector Classification". In: *IEEE Transactions on neural networks and learning systems* 28.5 (2017), pp. 1241–1248.

data and the classification is performed on the superpixels (rows of $\bar{\mathbf{F}}$).

2.2.3.2. KNN. Following the same previous introduced notation, denote $\mathcal{N}_{j_s}$ as the set containing the k_n nearest neighbors of the j_s pixel $\bar{\mathbf{f}}_{j_s}^T$ with labels denoted as $\{\ell_1, \ell_2, \dots, \ell_{k_n}\}$. Then, the KNN classifier finds the k_n nearest neighbors of the $\bar{\mathbf{f}}_{j_s}^T$ point in the training data and assigns such testing point to the most frequently occurring class of its k_n neighbors. The KNN perform the classification with the following majority voting rule:

$$\text{Class}(\bar{\mathbf{f}}_{j_s}^T) = \arg \max_{\Omega=\Omega_1, \dots, \Omega_c} \sum_{i=1}^{k_n} \delta(\ell_i, \Omega), \quad (2.13)$$

where $\delta(\cdot, \cdot)$ is the Kronecker delta function which is equal to 1 if all its arguments are equal, and 0 otherwise.

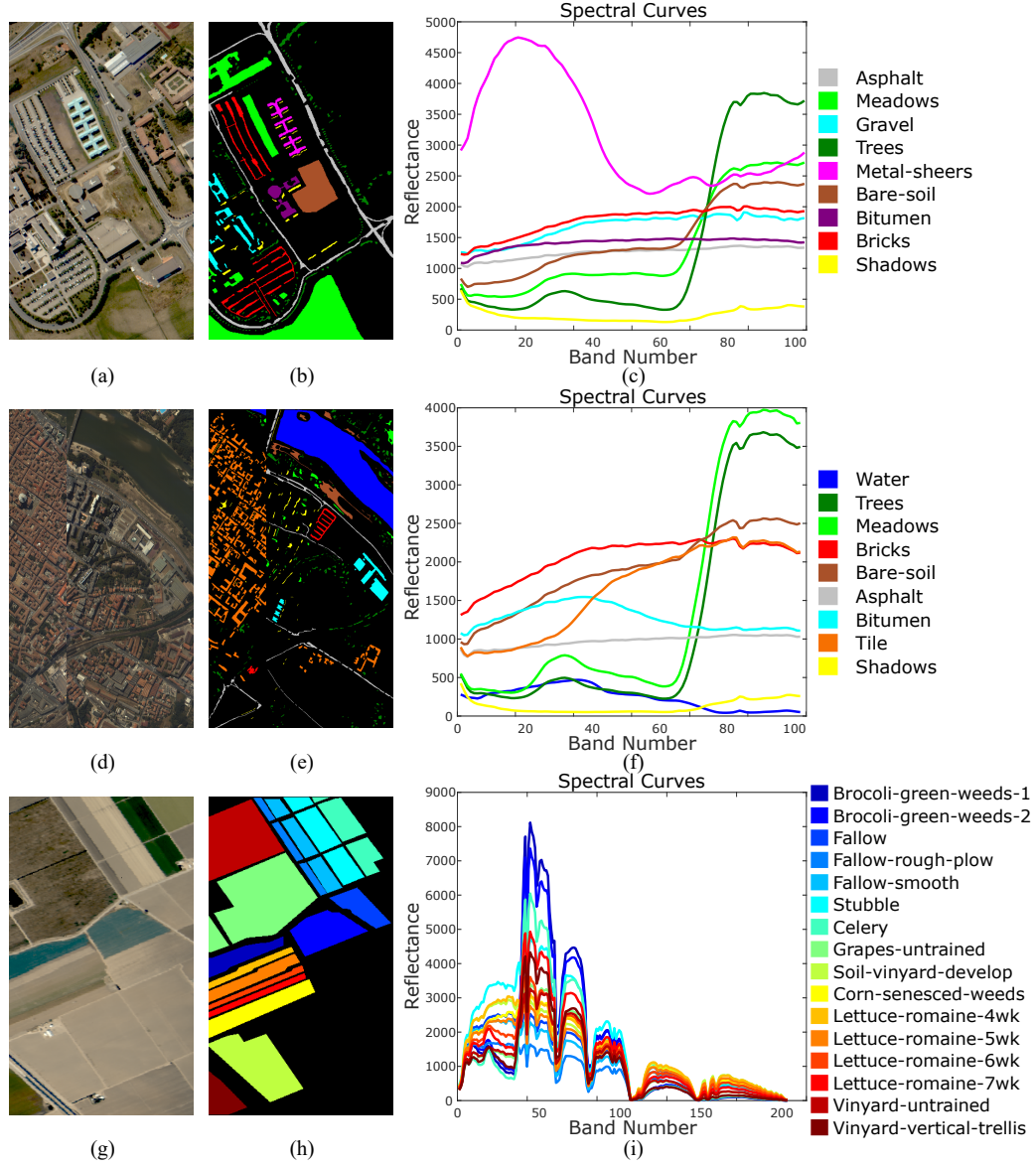
2.3. SIMULATIONS AND RESULTS

In this section, the performance of the proposed spectral image classification method is evaluated. In particular, we test the classification framework on three datasets, the Pavia University, the Pavia Center, and the Salinas Valley. The results are compared against other approaches that fuse high spatial and spectral resolution images, followed by the classification.

- **Pavia University** is an image acquired over an urban area surrounding the University of Pavia, northern Italy, by the Reflective Optics System Imaging Spectrometer (ROSIS-03) airborne sensor. The size of dataset is 610×340 pixels and 103 spectral bands, with a high spatial resolution of 1.3 meters per pixel and a spectral coverage ranging from 0.43 to $0.84 \mu\text{m}$. Figure 8 (a)-(c) shows the color composite of Pavia University, the corresponding ground truth image which differentiates nine land-cover classes: asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks and shadows; and the characteristic spectral signatures of each class, respectively.
- **Pavia Center** is an image also acquired with the ROSIS-03 sensor with size of 1096×715 pixels and 102 spectral bands, and it retains the same spatial and spectral resolution characteristics as the Pavia University dataset. The ground truth of Pavia Center contains nine classes: water, trees, asphalt, self-blocking bricks, bitumen, tiles, shadows, meadows and bare soil. A false-color image of the Pavia Center, the ground truth and its spectral signatures are shown in Fig.8 (d)-(f).
- **Salinas image.** This scene was collected in 1998 by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) on Salinas Valley, California, USA. The size of dataset is 512×217 pixels and 204 spectral bands in the range of 0.24 to $2.40 \mu\text{m}$. The Salinas ground truth contains 16 land-cover classes: broccoli green weeds 1, broccoli green weeds2, fallow, fallow rough plow, fallow smooth, stubble, celery, grapes untrained, soil vinyard develop, corn senesced green weeds, lettuce romaine 4wk, lettuce romaine 5wk, lettuce romaine 6wk, lettuce romaine 7wk, vinyard untrained, and vinyard vertical trellis. Figure 8 (g)-(i) shows a false-color of Salinas Valley, its ground truth and the spectral signatures of each class, respectively.

Following the matrix model described in section 2.1, the HS images ($\mathbf{F}_h$) are obtained with a decimation ratio across the spatial coordinates of $p = 4$. Therefore, the selected region

Figure 8: False-color images, ground truth maps and spectral signatures of each land-cover class for Pavia University (a)-(c), Pavia Center (d)-(f), and Salinas Valley (g)-(i), respectively.



of the Pavia University was resized to a high-resolution-spectral image of $152 \times 85 \times 103$; the Pavia Center image was spatially reduced to 274×178 pixels and 102 spectral bands; and the Salinas Valley image to $128 \times 54 \times 204$. On the other hand, with a selection matrix applied to the original datasets, a high-resolution-spatial RGB (F_m) image of each dataset was obtained. The size of the F_m images for Pavia University, Pavia Center and Salinas are $610 \times 340 \times 3$, $1096 \times 715 \times 3$, and $512 \times 217 \times 3$, respectively.

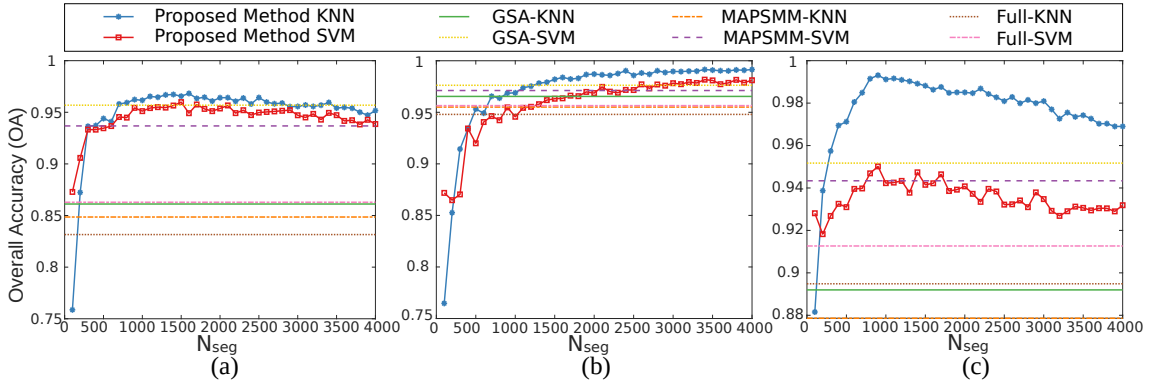
The experiments described below were performed by randomly selecting 5% of the labeled referenced data as training samples. Note that, in the presented results, we refer to "Full-data" when classifying the complete high spatial and spectral image, i.e., no fusion was performed. The aim of adding the Full-data results is to evidence the impact of incorporating spatial information in the classification performance. In addition, we compare the proposed approach with performing the fusion procedure with two state-of-the-art al-

gorithms, before classifying the obtained high spatio-spectral pixels. Specifically, "GSA" refers to the results achieved classifying a high spectral and spatial resolution image obtained using the state-of-the-art fusion method GS Adaptive (GSA) algorithm⁴⁰. Similarly, "MAP-SMM" corresponds to the results achieved classifying a high spectral and spatial resolution image reconstructed with the MAP-SMM algorithm⁵¹. In all the simulations, the datasets were classified with the SVM and KNN algorithms.

2.3.1. Impact of the number of superpixels and training samples in the classification accuracy. In the first experiment, the number of superpixels to cluster the image was varied to analyze the impact in the overall accuracy OA. In order to identify an adequate number of segments that provides the highest classification accuracy for each dataset, the value of N_{seg} was varied from 100 to 4000 in steps of 100. Figure 9 shows the overall accuracy results of the proposed method with both classifiers, SVM and KNN, by varying the number of segments N_{seg} in each dataset.

Figure 9 (a) presents the obtained results for the Pavia University. Note that, as the number of desired segments (N_{seg}) increases, the overall accuracy increases exponentially until reaching $N_{seg} = 1600$ where the OA begins to decrease slowly, as depicted with the red and blue lines. Furthermore, note that the proposed method with KNN outperforms the results obtained by classifying the full-data and data obtained by the GSA and MAP-SMM fusion methods when $N_{seg} \geq 700$ and $N_{seg} \leq 2800$.

Figure 9: Overall accuracy of classification varying the number of superpixels (N_{seg}) in the proposed method when classifying (a) Pavia University, (b) Pavia Center, and (c) Salinas Valley. The obtained OA when using the full-data, and the GSA and the MAP-SMM fusion methods are provided for comparison purposes.



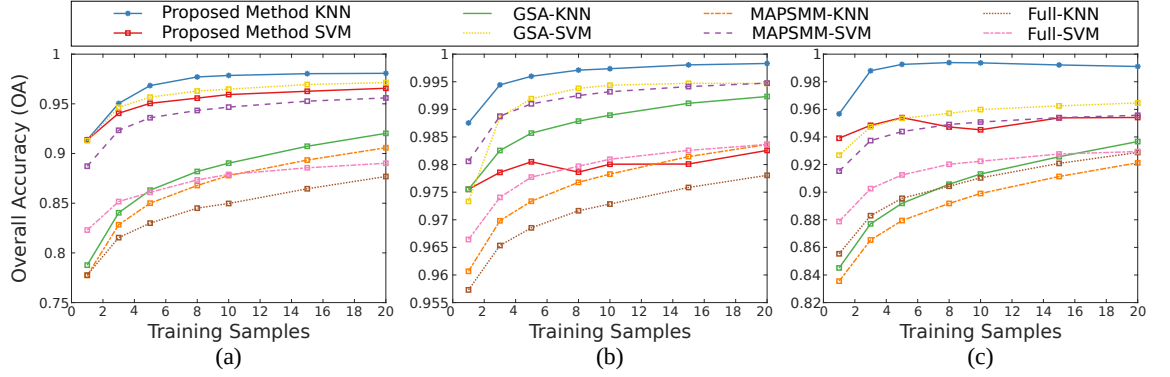
The obtained results for Pavia Center are shown in Fig. 9 (b). As observed, when the number of superpixels increases, the performance of the classification improves exponentially, until $N_{seg} \approx 3400$ where the overall accuracy begins to tend to $\approx 99.61\%$. As observed, in this particular image, all the methods perform well, however, with an adequate selected number of segments, the proposed method with KNN outperforms the others.

Figure 9 (c) shows the results obtained for the Salinas image. In this experiment, using the KNN classifier, the classification accuracy increases exponentially when the number of superpixels increases. However, when $N_{seg} = 900$, the OA begins to decrease. In this

⁵¹ Michael T Eismann and Russell C Hardie. "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.3 (2005), pp. 455–465.

experiment, it is also important to note that the performance of the KNN classifier was significantly better than SVM. In addition the results obtained with the proposed method and the KNN classifier outperforms all the other methods when $N_{seg} \geq 300$.

Figure 10: Overall accuracy of classification varying the number of training samples in the proposed method when classifying (a) Pavia University, (b) Pavia Center, and (c) Salinas Valley. The obtained OA when using the full-data, and the GSA and the MAP-SSM fusion methods are provided for comparison purposes.



In the second experiment, the number of samples used in the training stage was varied to analyze the impact in the OA. As observed in Fig. 10, all results obtained from the KNN classifier provides low-accuracy classification results in comparison with SVM, except for the proposed method which exhibits good performance for the tested number of training samples. This is expected since the KNN classifier takes more advantage of the spatial contextual information given by superpixels.

In general, for the proposed method with KNN, increasing the number of training samples will also improve the overall accuracy. However, when the number of training samples is greater than 8, the overall accuracy does not vary notably. In addition, from Fig. 4.(a), it can be noted that, for 1% of training samples, the proposed method with both KNN and SVM classifiers provides the same results as those given by GSA-SVM.

2.3.2. Classification Maps and Quantitative Results. To further validate the performance of the proposed framework, classification maps and quantitative results from Full data, GSA, MAP-SMM and the proposed approach are presented in Figures (4)-(6) and Tables (1)-(3).

Table 1 shows the numerical results for each in the nine land-cover classes in Pavia University. The presented results of average accuracy (AA), OA, the Kappa coefficient, and the running time are the average of 25 realizations of each experiment. In the table, the optimal values are shown in bold font and the second-best results are underlined. All the numerical results, except the Kappa coefficients and the running time, are given in percentage.

From Fig. 11 and Table 1, it can be clearly seen that the proposed classification approach provides the best accuracy results in comparison with classifying the pixels of the full and the fused spectral datacubes. The OA values for the different classification approaches are shown in each subfigure of Fig. 11. For this dataset, the best accuracy classification results are provided by the proposed framework using the SVM classifier. However, using the proposed method with the KNN classifier provides the best classification time. All execution times shown in this paper includes the time of the fusion stage.

Figure 12 and Table 2 show the visual and numerical results of the Pavia Center dataset. Notice that when the KNN classifier is used with the proposed approach, the best overall accuracy and kappa coefficient are obtained. In addition, this method takes the shortest execution time. On the other hand, note that although using the SVM classifier with the proposed method the classification results are lower than using KNN, these are the second-best results in terms of overall accuracy and time.

Table 1: Quantitative results of the different classification methods for the Pavia University image (OA%).

Class	FULL		GSA		MAP-SMM		PROPOSED	
	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN
Asphalt	86.41	76.48	97.04	81.00	94.94	82.40	96.10	<u>96.28</u>
Meadows	95.75	94.74	98.89	95.47	97.79	94.28	99.04	99.60
Gravel	69.18	62.34	87.34	66.72	82.93	70.83	<u>90.40</u>	96.61
Trees	91.51	83.74	94.07	85.96	91.91	74.66	74.39	77.81
Metal sheets	99.40	99.43	<u>99.87</u>	99.92	<u>99.00</u>	98.90	99.16	99.27
Bare soil	61.19	55.81	87.59	70.90	82.13	66.00	<u>92.46</u>	99.44
Bitumen	72.41	79.02	94.73	76.27	95.08	80.56	98.70	98.06
Bricks	76.53	77.89	93.03	76.58	91.60	79.32	<u>93.34</u>	93.39
Shadows	95.81	94.91	<u>97.56</u>	98.89	95.73	94.02	95.80	96.60
AA	83.13	80.48	<u>94.46</u>	83.52	92.34	82.33	93.27	95.23
OA	86.24	83.01	<u>95.69</u>	86.11	93.68	84.85	95.00	96.66
Kappa	0.8164	0.7731	<u>0.9429</u>	0.8158	0.9162	0.7985	0.9337	0.9559
Time	16.7052	23.9618	16.5859	25.1008	27.0535	38.1522	<u>10.8399</u>	6.8039

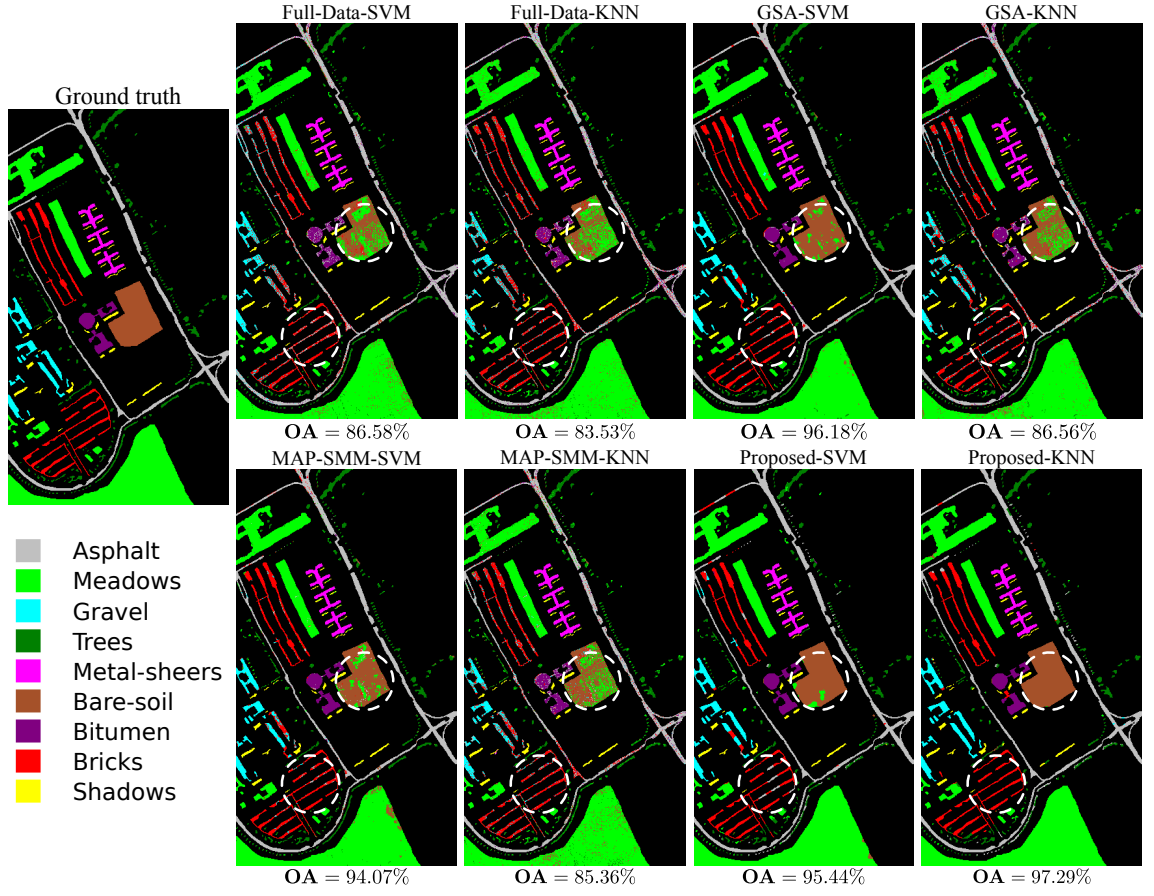
Table 2: Quantitative results of the different classification methods for the Pavia Center image (OA%).

Class	FULL		GSA		MAP-SMM		PROPOSED	
	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN
Water	99.97	99.97	99.98	<u>99.99</u>	99.98	99.98	99.96	100
Trees	93.75	89.98	<u>97.54</u>	<u>94.25</u>	95.15	93.41	97.27	98.56
Meadows	86.27	87.58	<u>95.95</u>	92.02	87.16	87.27	93.07	97.37
Bricks	77.20	74.01	92.45	89.97	91.01	82.88	99.48	<u>97.55</u>
Bare soil	95.94	94.87	97.23	97.50	97.34	92.92	99.38	<u>98.85</u>
Asphalt	91.69	86.24	<u>97.40</u>	94.34	95.39	90.30	97.13	99.29
Bitumen	92.54	90.46	97.59	93.89	87.43	92.02	99.07	98.56
Tile	99.63	99.15	99.67	<u>99.85</u>	99.58	99.20	99.57	99.93
Shadows	96.70	94.32	98.32	99.55	90.97	98.29	90.68	<u>99.10</u>
AA	92.63	90.73	<u>97.35</u>	95.71	93.78	92.92	97.29	98.80
OA	97.74	96.84	99.11	98.54	98.00	97.51	99.13	99.62
Kappa	0.9680	0.9552	0.9874	0.9794	0.9716	0.9647	<u>0.9877</u>	0.9946
Time	35.5553	180.1488	54.4082	359.6640	58.2847	353.3343	<u>10.9937</u>	7.3261

The performance of the classification of the Salinas Valley dataset is shown in Fig. 13 and the quantitative results are presented in table 3. This spectral image exhibits the greatest increment in classification performance when using the proposed method, compared to the other approaches. By applying the classifier directly on the original full data, OA classification results of 91.29% and 89.47% were obtained, in 33.424 and 17.903 seconds, respectively, whereas the proposed method using the KNN classifier achieves an OA of 99.21% in 1.414 seconds.

In addition, note that the second best classification results are obtained using the GSA fusion approach and the SVM classifier. However, the second best classification time was obtained by the proposed method and the SVM classifier. It is important to note that

Figure 11: Visual classification results on Pavia University image. Figure shows the ground truth, and the results for Full-Data, GSA, MAP-SMM, and the Proposed approach using both SVM and KNN classifiers.



the good classification results obtained by the GSA and the MAP-SMM are greatly due to the denoising process performed in those methods.

Figure 12: Visual classification results on Pavia Center image. Figure shows the ground truth, and the results for Full-Data, GSA, MAP-SMM, and the Proposed approach using both SVM and KNN classifiers.

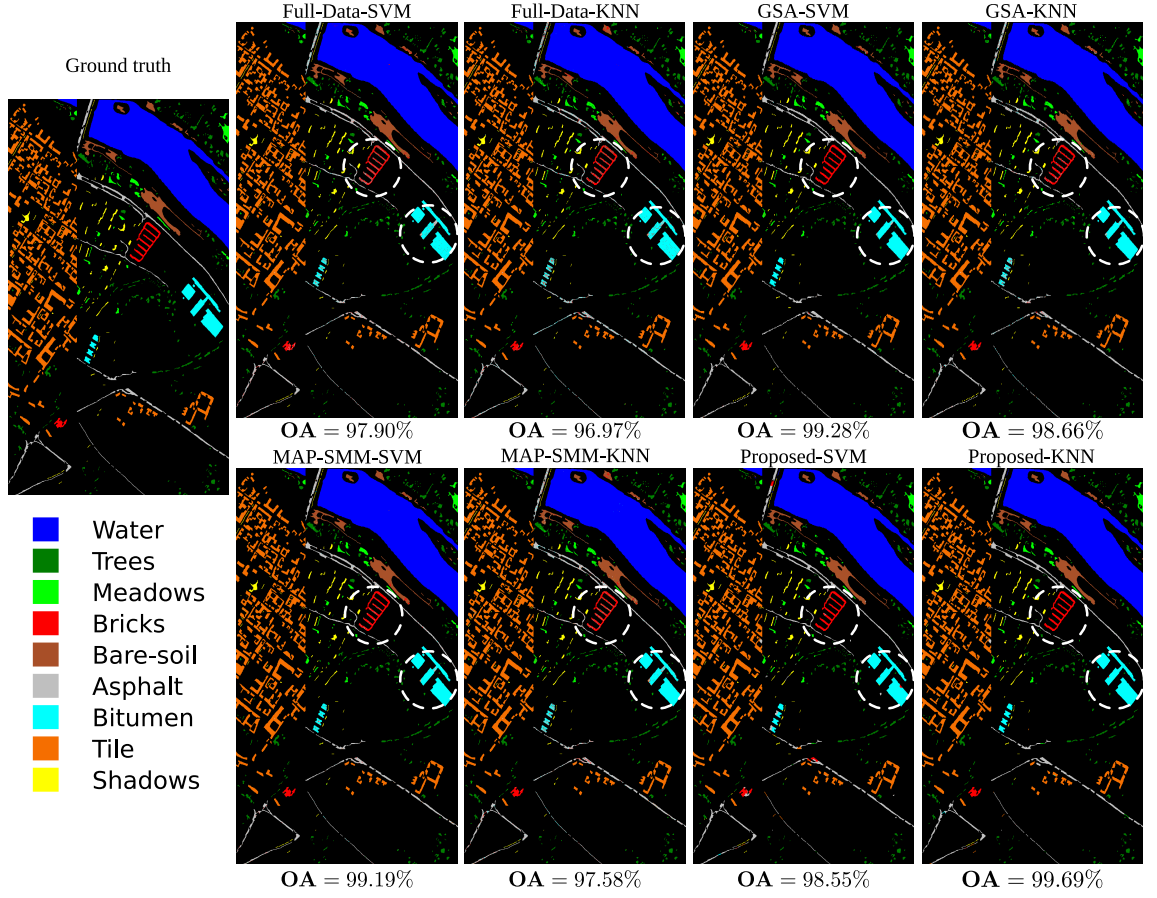
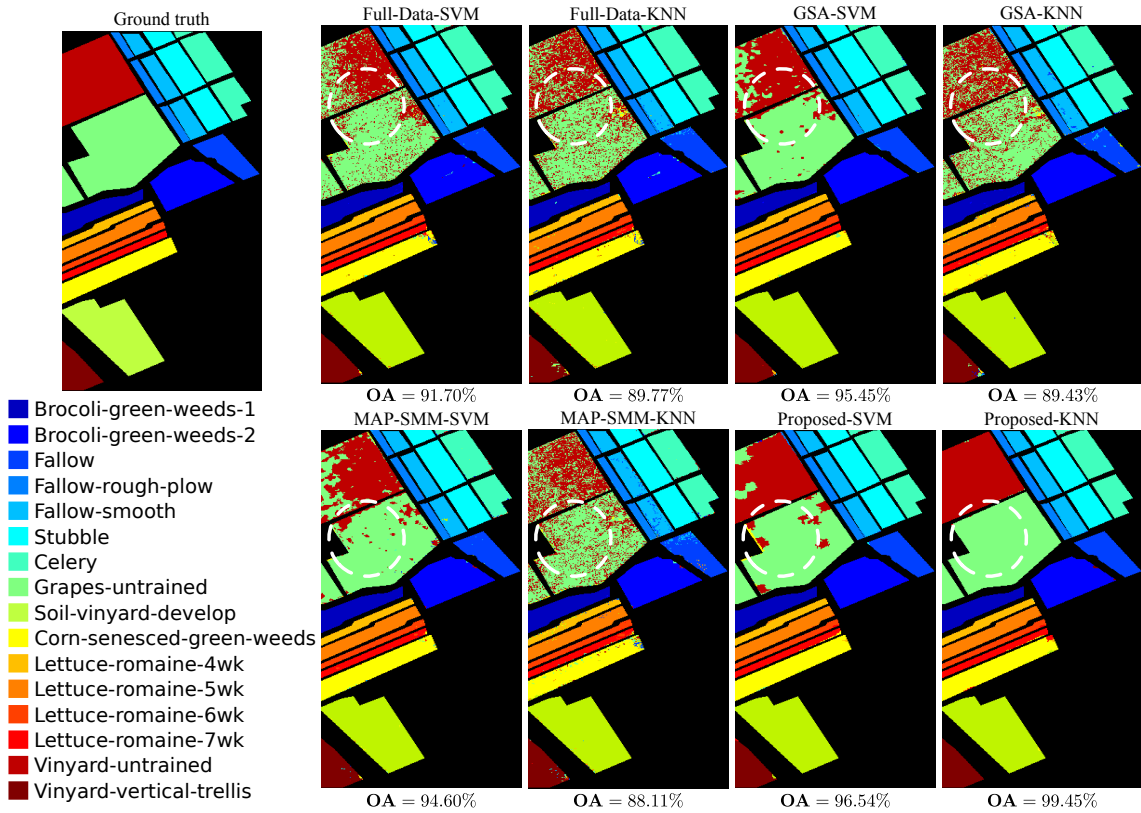


Table 3: Quantitative results of the different classification methods for the Salinas Valley image (OA%).

Class	FULL		GSA		MAPSMM		PROPOSED	
	SVM	KNN	SVM	KNN	SVM	KNN	SVM	KNN
Brocoli-green-1	98.83	99.00	99.48	98.13	99.47	99.74	100	99.86
Brocoli-green-2	99.23	99.10	99.91	99.78	99.33	99.12	99.80	99.60
Fallow	99.24	97.07	99.76	95.07	99.91	93.85	99.78	<u>99.78</u>
Fallow-rough-plow	98.86	99.52	98.09	98.39	98.58	98.87	99.65	<u>99.57</u>
Fallow-smooth	99.35	96.59	98.98	95.63	98.19	93.10	98.90	<u>98.95</u>
Stubble	99.83	99.70	99.99	99.86	99.86	99.50	99.92	99.79
Celery	99.47	99.41	99.48	98.96	99.08	98.61	99.75	99.60
Grapes-untrained	83.21	76.73	<u>92.60</u>	77.52	90.71	73.89	91.51	99.72
Soil-vinyard-develop	99.50	98.77	99.99	99.46	99.82	98.96	99.94	99.80
Corn-senesced-green-weeds	96.29	93.34	<u>98.32</u>	91.73	97.71	90.53	98.28	98.67
Lettuce-romaine-4wk	97.80	<u>98.25</u>	97.88	89.79	97.45	88.56	97.95	98.93
Lettuce-romaine-5wk	99.83	<u>99.65</u>	99.97	98.42	<u>99.90</u>	98.65	98.72	99.49
Lettuce-romaine-6wk	<u>98.01</u>	97.74	98.67	97.93	97.38	96.59	96.56	92.35
Lettuce-romaine-7wk	<u>96.63</u>	94.51	98.90	94.92	96.80	93.27	95.63	90.88
Vinyard-untrained	66.45	67.07	78.71	66.90	76.90	65.61	<u>78.98</u>	99.28
Vinyard-vertical-trellis	98.76	97.63	97.90	96.38	98.86	95.99	99.02	<u>98.88</u>
AA	95.71	94.63	<u>97.42</u>	93.68	96.87	92.80	97.15	98.45
OA	91.29	89.47	<u>95.17</u>	89.20	94.34	87.86	94.93	99.21
Kappa	0.9030	0.8828	<u>0.9462</u>	0.8797	0.9369	0.8649	0.9435	0.9912
Time	33.4255	17.9035	34.4387	18.8947	35.4682	18.4208	<u>5.4616</u>	1.4140

Figure 13: Visual classification results on Salinas Valley image. Figure shows the ground truth, and the results for Full-Data, GSA, MAP-SMM, and the Proposed approach using both SVM and KNN classifiers.



3. SUPERVISED CLASSIFICATION OF SPECTRAL IMAGES ACQUIRED WITH SINGLE PIXEL CAMERA AND SIDE INFORMATION

The proposed methodology in this project included the following stages:

- To design a mathematical model that describes the acquisition architecture of the measurements, the fusion model, and the supervised classification.
- To design and develop an algorithm which summarize the proposed mathematical model.
- To perform simulations with the designed algorithm, based on real hyperspectral data available on the Internet, and compare the results with the classification of the same data without taking into account side information.
- To implement in the optical laboratory of the HDSP research group an architecture for the simultaneous acquisition of SPC and RGB information. Also, to acquire data of cocoa beans with the testbed and to test the performance of the algorithm with these datasets.
- Finally, to prepare a document with the complete report of the work developed, including the analysis and evaluation of the results.

In this section it is presented the result of the first two stages of the methodology: mathematical model and algorithm.

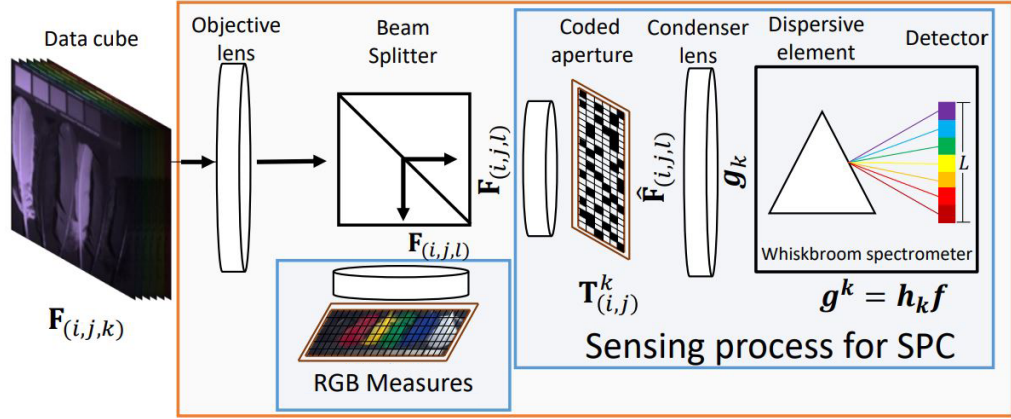
3.1. MATHEMATICAL SENSING MODEL

Compressive spectral imaging (CSI) architectures allow reconstructing spectral images from a lower number of measures than the traditional scanning-based methods. In these architectures, the coded aperture design is critical to obtain high-quality reconstructions. The structure of coded apertures is traditionally designed without information about the scene, but recently side information-based architectures provide prior details of the stage, which enables adaptive coded aperture designs. With the development of the work shown in section 2, it was evident that incorporating the side spatial information to the spectral measurements using superpixels allows improving the time and accuracy of the classification. For these reasons, this section proposes the development of an adaptive coded aperture design for spectral imaging with the single pixel camera, based on a multi-resolution approach.

3.1.1 Side information-based SPC sensing model for spectral imaging. As illustrated in Fig. 14, the proposed acquisition system is composed of two arms, the first corresponds to the SPC architecture and the other acquires the side information with an RGB sensor.

3.1.1.1 Single pixel camera sensing model. The single pixel architecture illustrated in Fig. 14 employs a coded aperture $\mathbf{T}_{(i,j)}^k$ that spatially modulates all the spectral bands from the input scene $\mathbf{F}_{(i,j,l)}$ with the same pattern, where i, j index the spatial coordinates, l accounts for the spectral band, and k indexes each captured snapshot.

Figure 14: Scheme of the SPC used to acquire the spatial and spectral information with RGB side information.



Specifically, the coded aperture $\mathbf{T}_{(i,j)}^k$ is a binary pattern whose spatial distribution determines the quality of the reconstructed scene. In this case, the binary levels are either 1 or -1 . In practical terms, the modulation effect caused by the -1 entries can be implemented by acquiring a measurement with an all ones coded aperture and subtract it from each captured snapshot. Mathematically, the coded aperture effect over the input scene is represented as

$$\hat{\mathbf{F}}_{(i,j,l)}^k = \mathbf{F}_{(i,j,l)} \mathbf{T}_{(i,j)}^k, \quad (3.1)$$

Then, the modulated scene $\hat{\mathbf{F}}$ is concentrated in a single spatial point by the condenser lens, and captured by a whisk-broom spectrometer. This detector splits the incoming light rays in a discrete measure of the spectrum for each band as

$$g_l^k = \sum_i \sum_j \mathbf{F}_{(i,j,l)} \mathbf{F}_{(i,j)}^k, \quad (3.2)$$

for $i = 1, 2, \dots, M, j = 1, 2, \dots, N$, and $l = 1, 2, \dots, L$. The acquisition system can also be modeled as

$$g_l^k = \mathbf{h}_k^T \mathbf{f}_l, \quad (3.3)$$

where $\mathbf{h}_k$ is the vectorization of the coded aperture for each snapshot k , $\mathbf{f}_l$ is the vectorization of the l -th band of the SI $\mathbf{F}$. In general, the sensing model for all shots captured for the l -th band can be written as

$$\mathbf{g}_l = \mathbf{H} \mathbf{F}_l, \quad (3.4)$$

with $\mathbf{g}_l = [g_l^1, \dots, g_l^K]^T$, $\mathbf{H}$ is a $K \times MN$ matrix given by $\mathbf{H} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]$, and K is the number of shots. It is important to remark that each shot employs a different coded aperture pattern. Furthermore, the sensing model for all the spectral bands and K shots is given by

$$\mathbf{g} = \hat{\mathbf{H}} \mathbf{F}, \quad (3.5)$$

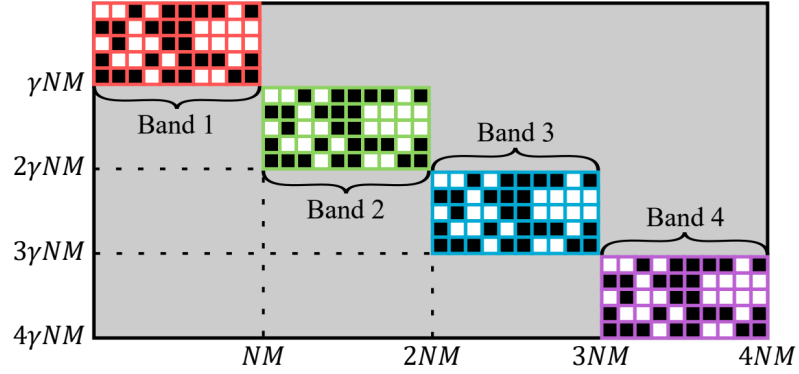
where $\mathbf{g} = [(\mathbf{g}_0)^T, \dots, (\mathbf{g}_{L-1})^T]^T$, $\hat{\mathbf{H}}$ is the sensing matrix whose structure is illustrated

in Fig. 15, which can be obtained as a block diagonal matrix

$$\hat{\mathbf{H}} = \mathbf{I}_L \otimes \mathbf{H}, \quad (3.6)$$

where $\mathbf{I}_L$ is an $L \times L$ identity matrix, such that the number of columns and rows of $\hat{\mathbf{H}}$ is MNL and γMNL , respectively, with $\gamma = \frac{K}{MN}$ as the compression ratio, taking values $\gamma \in [0, 1]$.

Figure 15: SPC sensing matrix $\hat{\mathbf{H}}$, with $M = 2$, $N = 5$, $\gamma = 0.5$ and $L = 4$. White squares represent a positive value 1 according to the coded aperture design, black squares represent -1 , and gray zones are 0.

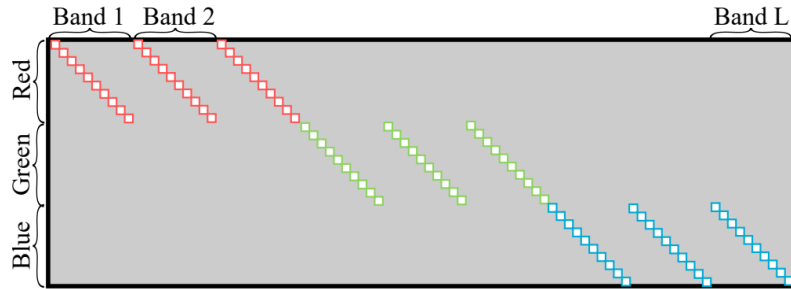


3.1.1.2 RGB Camera sensing model for spectral imaging. The model used for acquiring the RGB image assumes that the detector has an equal response for all spectral bands in each spectral channel (Red, green and blue). Therefore, the acquisition of the RGB image can be modeled as the linear system given by

$$\bar{\mathbf{g}} = \bar{\mathbf{H}}\mathbf{f}, \quad (3.7)$$

where $\bar{\mathbf{H}} = \mathbf{I}_3 \otimes (\mathbf{1}_{1 \times L/3} \otimes \mathbf{I}_{MN})$. Figure 16 shows an example of the structure of $\bar{\mathbf{H}}$, where white squares represent a 1 value for the color indicated by the border, and gray zones are zero. The acquired RGB image has an important role for this work, because it is used to design the coded aperture patterns. Furthermore, it can be also included in the reconstruction problem to improve the quality of the recovered data cube.

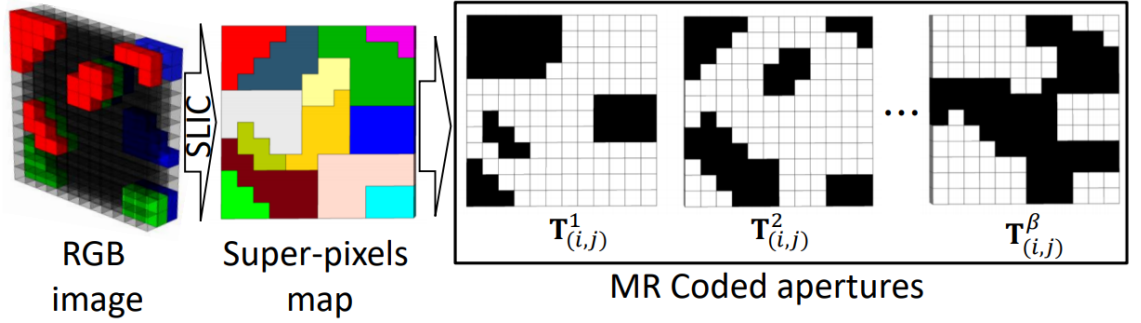
Figure 16: RGB sensing matrix $\bar{\mathbf{H}}$, for $M = 2$, $N = 5$, and $L = 9$. White squares represent 1 and gray zones are 0.



3.1.2 Coded aperture design based on Super-pixel. The input for the coded aperture design is the RGB image captured as in (3.7). Specifically, it is used to identify the

uniform zones in the scene, such that similar pixels can be grouped into a super-pixel. The resulting map of super-pixels is then used to design the coded aperture pattern. The main goal is to obtain multi-resolution (MR) coded apertures whose features match with the super-pixel map. First, to determine the super-pixel map, a clustering method such as the SLIC algorithm⁵² is employed. A simple example of this process is illustrated in Fig. 17, where each color represents a super-pixel. Note that the color super-pixel map is decomposed as β coded aperture patterns, where β is the desired number of super-pixels, and it is a required input for the SLIC method. To obtain the set of CA $\mathbf{T}$, the map of super-pixels is used to generate a multi-resolution decimation matrix Δ , as in⁵², where each row of Δ represents the integration of pixels into each super-pixel. It is worth noting that the size of Δ depends on the β as $\Delta \in \mathbb{R}^{\beta \times MN}$. Moreover, the coded aperture design is directly related to the design of the matrix $\mathbf{H}$ from (3.4) which, in turn, determines the sensing matrix $\hat{\mathbf{H}}$ in (3.6).

Figure 17: Scheme of the process to generate MR coded apertures. Each superpixel of the RGB scene is associated to a single coded aperture block.



Several SPC snapshots are required in order to provide good quality reconstructions. In this work, we propose to capture β measurements, in other words, the number of snapshots is equal to the number of desired super-pixels. This condition imposes that

$$\text{rank}(\mathbf{H}) = \beta. \quad (3.8)$$

Therefore, a full rank matrix $\mathbf{W}$ is used to design the sensing matrix for each spectral band as

$$\mathbf{H} = \mathbf{W}\Delta. \quad (3.9)$$

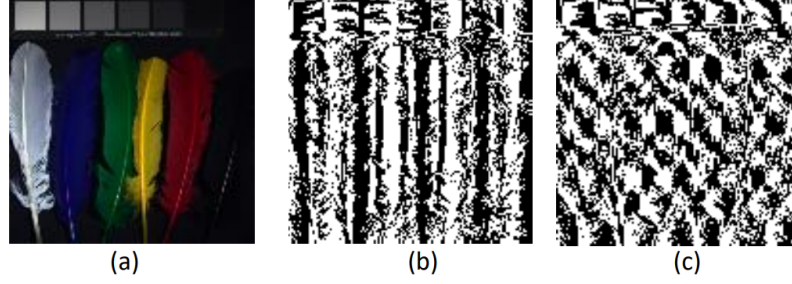
Specifically, $\mathbf{W} \in \mathbb{R}^{\beta \times \beta}$ is a Hadamard matrix, because its inverse is the same transposed escalated matrix. This property enables the fast multi-resolution reconstruction approach proposed in Section 3.1.5.

Coded apertures for each snapshot are obtained by rearranging each row of $\mathbf{H}$ in (9) as a matrix. The resulting MR coded apertures, whose block features match the super-pixels from Δ are also illustrated in Fig. 17.

Furthermore Fig. 18(b) and (c) show two MR coded apertures obtained with the proposed design for the scene in Fig. 18(a).

⁵² Hans Garcia et al. "Multi-resolution reconstruction algorithm for compressive single pixel spectral imaging". In: *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 468–472.

Figure 18: Examples of the designed MR coded apertures for the spectral image in (a), using the proposed approach on (3.9).



3.1.3 Reconstruction with Side Information. In order to reconstruct the spatio-spectral data cube, the SPC measurements from (3.5) are stacked with those from the RGB camera in (3.7), as well as the corresponding sensing matrices

$$\tilde{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{H}} \\ \bar{\mathbf{H}} \end{bmatrix}, \tilde{\mathbf{g}} = \begin{bmatrix} \mathbf{g} \\ \bar{\mathbf{g}} \end{bmatrix}. \quad (3.10)$$

Note that $\hat{\mathbf{H}}$ is obtained as in (3.6), using the matrix $\mathbf{H}$ from (9).

Then, the spatio-spectral data cube can be obtained by solving the regularization problem given by

$$\tilde{\mathbf{f}} = \Psi \left\{ \underset{\tilde{\Theta}}{\operatorname{argmin}} \|\tilde{\mathbf{H}}\Psi\tilde{\Theta} - \tilde{\mathbf{g}}\|_2^2 + \tau\|\tilde{\Theta}\|_1 \right\}, \quad (3.11)$$

where τ is a regularization parameter, $\tilde{\Theta}$ is the sparse representation of $\tilde{\mathbf{f}}$ in the basis Ψ . It is important to remark that the new compression ratio $\bar{\gamma}$, taking into account the RGB image is given by

$$\bar{\gamma} = \frac{\gamma(L+3)}{L}, \quad (3.12)$$

where $\bar{\gamma} \approx \gamma$ as L increases.

3.1.4 Fast Multi-Resolution (FMR) Reconstruction. Several applications require fast SI reconstructions, however, traditional reconstruction methods solve the $l_2 - l_1$ norm optimization in (3.11) whose complexity depends on the data dimensions. To date, fast reconstruction alternatives have been developed⁵³. In this work, we propose to obtain a fast multi-resolution reconstruction based on the coded aperture design from section 3.1.2. FMR reconstructions of the SI can be obtained without solving a minimization problem, this is possible by taking advantage of the designed coded aperture with a Hadamard structure. Using this approach, the complexity to obtain the SI is reduced only to one matrix product. First, taking into account that for some Hadamard matrix $\mathbf{W} \in \mathbb{R}^{\beta \times \beta}$.

$$\mathbf{W}^T \mathbf{W} = \beta \mathbf{I}_\beta, \quad (3.13)$$

where $\mathbf{I}_\beta$ is a $\beta \times \beta$ identity matrix. If we multiply (3.4) by $\mathbf{W}^T$, and replacing $\mathbf{H}$ as in (3.9) we obtain that

$$\bar{\mathbf{f}}_l = (1/\beta) \mathbf{W}^T \mathbf{g}_l = (1/\beta) \mathbf{W}^T \mathbf{W} \Delta \mathbf{f}_l. \quad (3.14)$$

⁵³ Henry Arguello, Claudia V Correa, and Gonzalo R Arce. "Fast lapped block reconstructions in compressive spectral imaging". In: *Applied Optics* 52.10 (2013), pp. D32–D45.

Then, multiplying $\bar{\mathbf{f}}$ by some matrix $\hat{\Delta}$, such that $\hat{\Delta}\Delta \approx \mathbf{I}$ yields

$$\tilde{\mathbf{f}}_l = \hat{\Delta}\bar{\mathbf{f}}_l = (1/\beta) \hat{\Delta}\mathbf{W}^T\mathbf{W}\Delta\mathbf{f}_l \approx \mathbf{f}_l. \quad (3.15)$$

Thus, the best option is to have $\hat{\Delta}$ as the transpose of Δ , with normalized rows. Then, a FMR reconstruction $\tilde{\mathbf{f}}$ of the SI is obtained with high quality at low computational complexity, when $\mathbf{W}$ is a Hadamard matrix and $\mathbf{H}$ follows the design in (3.9).

3.2. SUPERVISED CLASSIFICATION MODEL FOR SPC MEASUREMENTS AND SIDE INFORMATION

Once reconstructed the underlying spatio-spectral scene, the traditional procedure to perform spectral signature classification is to determine the class labels of the image by training a support vector machine classifier (SVM) using all the pixels. However, taking advantage of the proposed coded aperture design described in section 3.1.2, it is possible to design a strategy that focuses on directly classify the SPC measurements, not only boosting the computational performance but also improving the classification accuracy.

Recalling Eq. 3.13 and 3.14 we have that

$$\bar{\mathbf{f}}_l = (1/\beta)\mathbf{W}^T\mathbf{W}\Delta\mathbf{f}_l = \Delta\mathbf{f}_l \quad (3.16)$$

for each spectral band l , where $\mathbf{f}_l$ is a vectorized spectral band. Note that, the product $\Delta\mathbf{f}_l$ decimates the $\mathbf{f}_l$ vector by integrating similar pixels into a single super-pixel. It can be seen as a feature extraction preprocessing step for classification, which takes into account the pixel neighboring information. In this sense, denoting

$$\bar{\mathbf{F}} = [\bar{\mathbf{f}}_1, \dots, \bar{\mathbf{f}}_L] \in \mathbb{R}^{\beta \times L}, \quad (3.17)$$

as the feature matrix which contains the spectral information of each super-pixel, then it is only necessary to determine the class label of the β spectral super-pixels (rows of $\bar{\mathbf{F}}$). Then, the pixels in $\mathbf{F}$ acquire the same label of the super-pixel, to which they belong.

Algorithm 1 summarizes the proposed model that takes into account the matrices $\hat{\Delta}$, Δ and $\bar{\mathbf{F}}$ along with an SVM classifier to perform the classification. Notice that, as described in section 2.2, the SVM training is performed using the rows of $\bar{\mathbf{F}}$ as training data and the classification is performed on the super-pixels (rows of $\bar{\mathbf{F}}$). The proposed method has mainly three advantages:

- The classification is performed directly on the SPC measurements $\mathbf{g}$, avoiding the expensive signal reconstruction which involves solving the problem in Eq. 3.11.
- As the number of super-pixels $\beta \ll MN$, the proposed approach boosts the computational performance.
- Due to the use of super-pixels, the neighboring spatial information is taken into account to extract the classification features; hence improving the accuracy.

These advantages are showed in the next Section.

Algorithm 1 SPC Classification Model using super-pixels

Input: $\mathbf{f}, \bar{\mathbf{H}}, \beta, c, v$.

Super-pixels Extraction :

- 1: $\bar{\mathbf{g}} = \bar{\mathbf{H}}\mathbf{f}$. ▷ Obtain the RGB image
- 2: $\hat{\Delta}, \Delta \leftarrow \text{SLIC}(\bar{\mathbf{g}}, \beta)$ ▷ should satisfies $\hat{\Delta}\Delta \approx \mathbf{I}$

Coded Aperture Desing :

- 3: $\mathbf{H} = \mathbf{W}\Delta$ ▷ Where $\mathbf{W}$ is a $\mathbb{R}^{\beta \times \beta}$ Hadamard matrix

SPC Measurement Acquisition :

- 4: $\mathbf{g}_l = \mathbf{H}\mathbf{F}_l$ ▷ Each column of $\mathbf{F}_l$ is a vectorized spectral band $\mathbf{f}_l$

Feature Extraction and Classification

- 5: $\bar{\mathbf{F}} = [\Delta\mathbf{f}_1, \dots, \Delta\mathbf{f}_L] = [\bar{\mathbf{f}}_1, \dots, \bar{\mathbf{f}}_L]$
 - 6: $\Omega = \{\omega_1, \dots, \omega_c\}$ ▷ Denote the set of c class labels
 - 7: $\Theta = \{\theta_1, \dots, \theta_v\}$ ▷ Denote the set of v integers indexing the training samples
 - 8: Train a SVM classifier solving the optimization problem in Eq. 2.11.
 - 9: Classify the test pixels solving Eq. 2.12.
-

4. EXPERIMENTS AND RESULTS

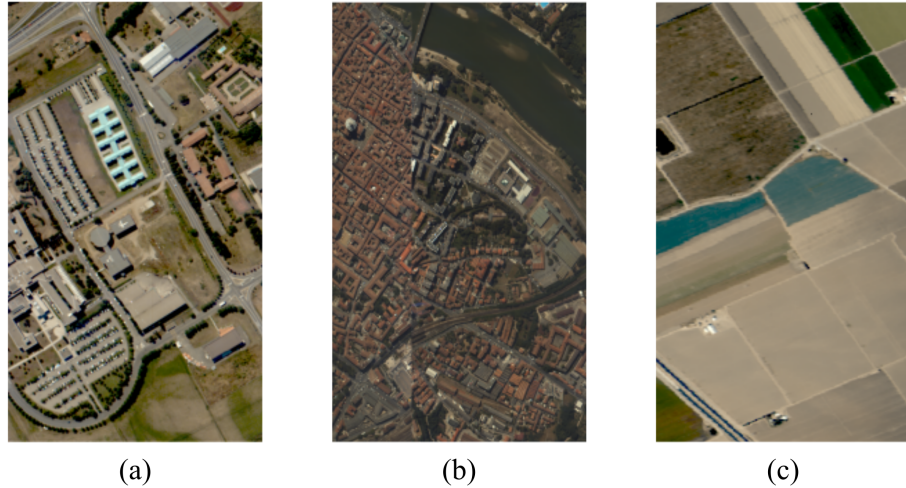
In this section, are carried out the final stages of the methodology proposed for the development of the work *"A supervised classification system of spectral images acquired with a single pixel optical architecture and side information"*.

First, several simulations were realized to test the performance of the proposed approach using traditional hyperspectral datasets available on the web. Next, is shown the optical architecture implemented in the laboratory for the simultaneous acquisition of spatial-spectral information with a single pixel camera and side information of an RGB sensor. Then, are presented the datasets acquired in the laboratory. Finally, are discussed the classification results obtained with the proposed method for data acquired experimentally in the laboratory. The results analysis include the study of the effect of the super pixels number, and therefore, the compression level of the measurements in the classification performance; the accuracy of the classification system by adding different noise levels; and, the visual maps and quantitative results of the classification.

4.1. SIMULATIONS

The proposed spectral images classification approach was tested on three real hyperspectral datasets, with different imaging environments, acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and the Reflective Optics System Imaging Spectrometer (ROSIS). Specifically, are used the Pavia University, the Pavia Center, and the Salinas Valley dataset, i.e., the same spectral images used in section 2. Fig. 19 shows the false-color composite of the images used.

Figure 19: False-color composite of (a) Pavia University, (b) Pavia Center, and (c) Salinas Valley.

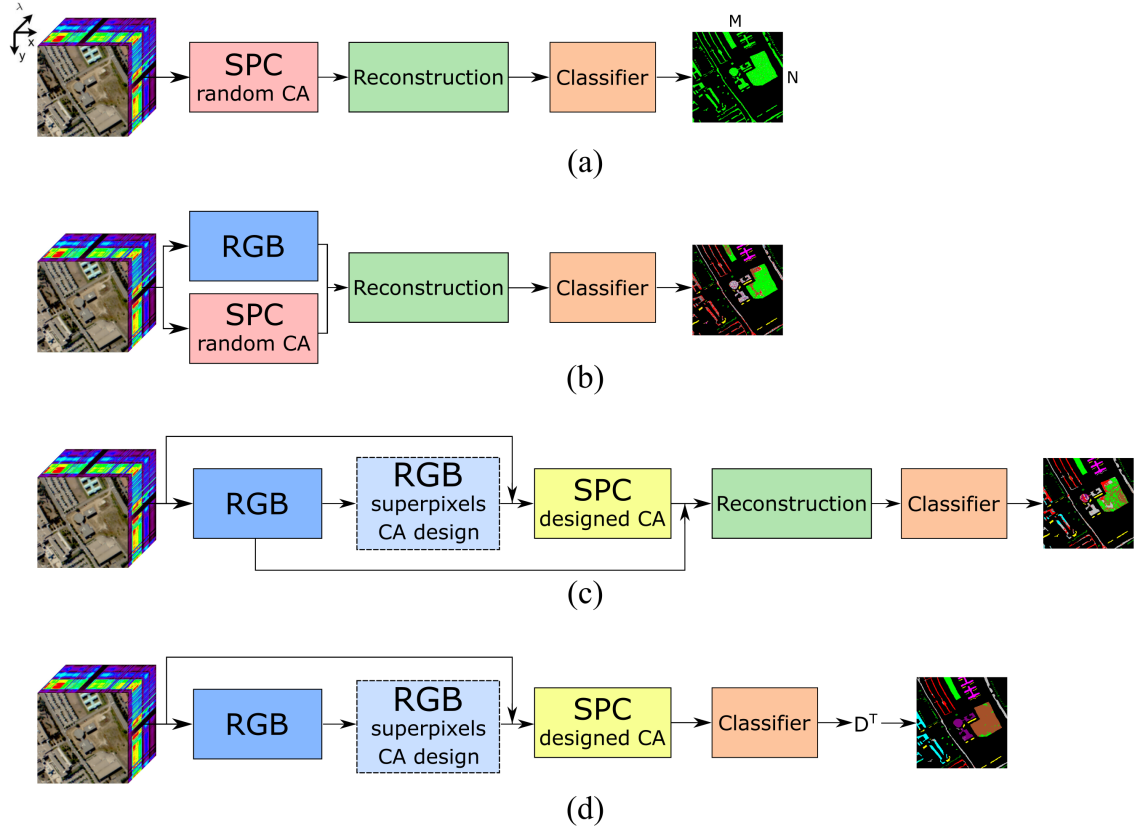


The performance of the proposed method is tested: Supervised classification of compressed SPC measurements acquired with aperture codes designed from the segmentation into superpixels of an RGB image of the same scene to be classified. Additionally, simulations of three additional approaches are performed for comparison purposes. In the first approach is classified the reconstruction of the acquired measurements with the classical SPC using a traditional coded aperture (CA) based on randomly permuted rows of the Hadamard transform with unit-norm columns, i.e., the RGB image is not added to the SPC measurements (20 (a)). The second approach uses the proposed acquisition

architecture stacking the RGB image to the SPC measures (SPC + RGB) as in (3.10) with traditional Hadamard based patterns (20 (b)). The third approach uses the coded aperture design proposed in section 3 and the RGB image is stacked to the SPC measures as in (3.10) (RGB + SPC + CA design). These three approaches obtain the data cube reconstructions by solving (3.11) using the corresponding matrices, and then, the classifier is executed over the high-spatial-spectral reconstructed image. The fourth approach (Proposed) uses the same configuration as the third approach but the classifier is executed over the fusion measurements (3.1.4). This method does not stack the RGB image or reconstruct the merged measures to apply the classifier. Then, an upsampling is made to the label matrix of the superpixels to obtain the classification map.

Simulations are executed with the original spatial resolution of the datasets: Pavia University size 610×340 pixels, Pavia Center size 1096×715 , and Salinas Valley size 512×217 . The number of spectral bands is explained in each experiment.

Figure 20: Diagram of tested approaches: (a) SPC with random CA, (b) SPC with random CA + RGB, (c) SPC + RGB + CA design, and (d) Proposed method.



4.1.1 Impact of the Number of superpixels in the Classification accuracy. In the first experiment, the number of superpixels to cluster the RGB image was varied to analyze the impact in the OA. Note that due to the superpixels extracted from the RGB image are used for the design of the aperture code with which the SPC measurements are acquired, the number of segments is directly related to the compression ratio (Fig. 17). The experiment was executed for values of $N_{seg} = 2^n$ with n from 4 to 12 in steps of 1, i.e., for $N_{seg} = 16, 32, 64, 128, 256, 512, 1024, 2048$, and 4096. Furthermore, this experiment allows to identify the appropriate number of segments to obtain the highest classification accuracy in each dataset.

Figure 21 shows the overall accuracy results of the proposed method by varying the number of segments N_{seg} in the simulation of each dataset. For comparison purposes, the OA obtained by varying the superpixels in SPC, SPC + RGB, and SPC + RGB + CA design are also provided.

Figure 21: Overall accuracy of the classification by varying the number of superpixels (N_{seg}) in (a) Pavia University, (b) Pavia Center, and (c) Salinas Valley.

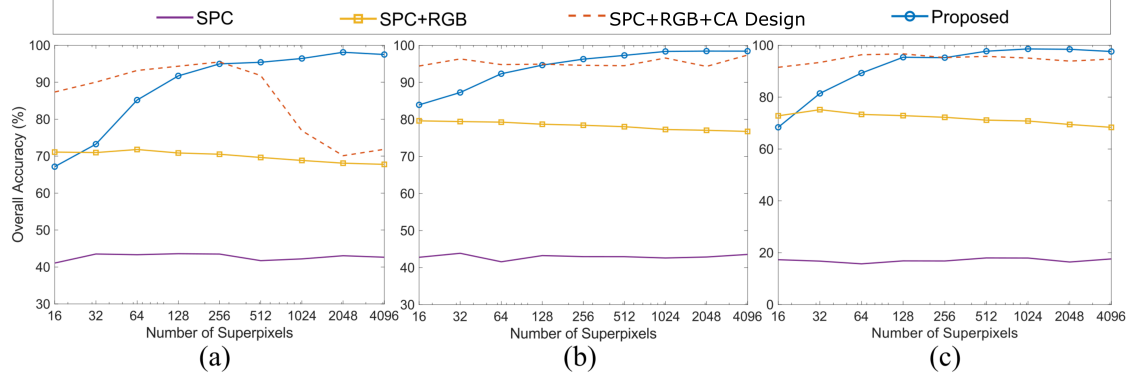


Figure 21 (a) presents the obtained results for the Pavia University. Note that, when the number of superpixels increases, the performance of the classification improves, until $N_{seg} \approx 2048$ where the overall accuracy begins to tend to $\approx 98.61\%$, as depicted with the blue line. Furthermore, note that the classification of SPC with the proposed coded aperture design + RGB, improves exponentially and outperforms the Proposed method from 256 superpixels (red line). In general, the two methods that use designed aperture codes outperform the two approaches that use random ones.

The obtained results for Pavia Center are shown in Fig. 21 (b). As observed, once more the performance of the Proposed method improves with increasing the number of superpixels, which from 512 segments tends to stabilize around $\approx 99.0\%$. On the other hand, the Proposed outperforms the SPC + RGB + CA design method from 128 superpixels. For this dataset, the three other approaches show a stable behavior as the number of segments varies. In the two approaches that use random aperture codes, this trend is expected since this parameter does not affect the rate compression.

Figure 21 (c) shows the results for the Salinas image dataset. In this experiment, a behavior similar to Pavia Center is obtained. The OA improves as the N_{seg} increases and from 512 superpixels tend to $\approx 99.6\%$. Note that, in the three simulations is evident the low accuracy obtained by SPC and the fusion approach that does not include the RGB measurements as prior information for the acquisition of SPC.

4.1.2 Visual maps and Quantitative results. In order to validate the classification performance of the proposed method, visual and quantitative results are presented in Figures (4.4)-(4.6) and Tables (4.1)-(4.3). In all the results shown below, the input parameter N_{seg} equals 512.

Figure 22 presents the obtained visual classification results on Pavia University. Table 4 shows the numerical results for each in the nine land-cover classes in this dataset. The presented results of average accuracy (AA), overall accuracy (OA), the Kappa coefficient, and the running time are the average of 5 realizations of each experiment. In the tables, the optimal values are shown in bold font and the second-best results are underlined. All

the numerical results are given in percentage, except the running time in seconds, and the Kappa coefficients.

From Fig. 22 and Table 4, it can be clearly seen that the proposed classification approach provides more accurate results than those obtained with the classification of SPC measurements without side information, and the classification of SPC with randomly aperture code stacked with the RGB image. The results obtained with the third approach also provide good overall accuracy results. However, the total time of the process through this approach is much greater than through the fourth, due to the additional time of the reconstruction stage.

Similarly, the visual classification results and quantitative evaluation on the Pavia Center are shown in Fig. 23 and Table 5. Table 5 shows a considerable increase in the classification accuracy with the proposed method compared to the other three approaches.

Finally, Fig. 24 and Table 6 presents the visual and numerical results on the Salinas Valley. In the test executed on this scene, a high similarity between the classification accuracy achieved with the approaches 3 and 4 is seen. Even, although the highest OA is attained by the proposed method, in this dataset the best AA is with the method 3.

Note that the proposed method achieves classification accuracy results superior to those obtained with the other approaches, even though it does not add the RGB image to the SPC measurements acquired with the codes designed to extract the classification features as it happens in the approaches 2 and 3. In contrast, the side information extracted as superpixels of the RGB image is optimally used to design the coding of the acquisition of SPC measurements, to the extent that the results are comparable (Table 6) or superior (Tables 4 and 5) to the other tested approaches.

In addition, in all datasets the proposed method takes significantly less total execution time than the others. The behavior of time is expected, mainly due to the absence of the reconstruction step (Reconstruction Time in Tables). Also, the Classification Time is less with this framework, because so many spectral signatures are classified as superpixels (N_{seg}), different from the other approaches whose number of signatures to be classified is equal to the number of pixels (MN) and $N_{seg} \ll MN$.

Figure 22: Visual classification results on Pavia University dataset

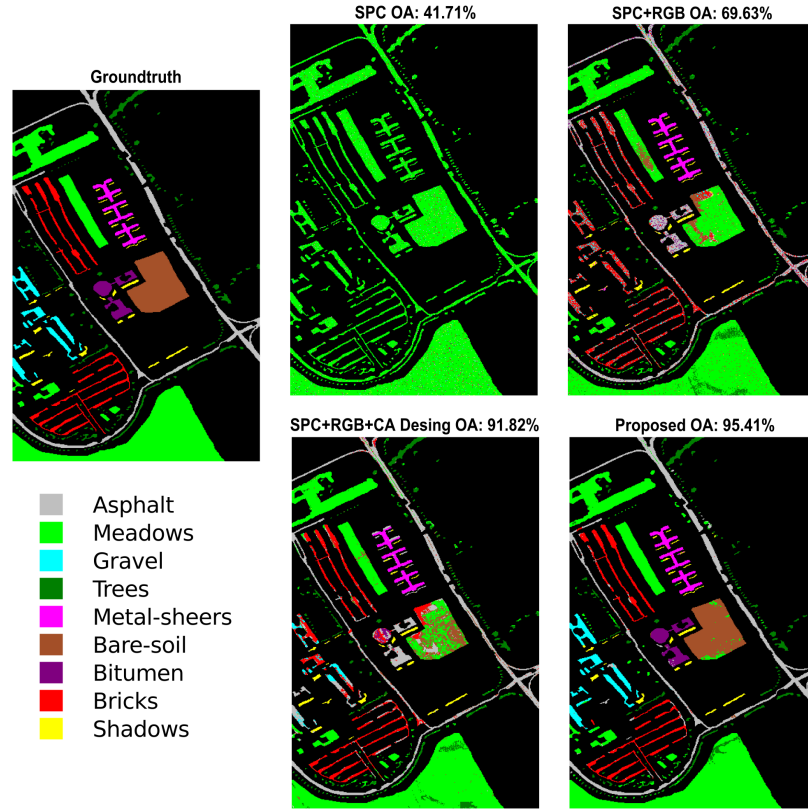


Table 4: Quantitative results on Pavia University dataset (OA %)

Class	SPC	SPC+RGB	SPC+RGB+CA Design	Proposed
Asphalt	2.35	74.56	90.79	97.50
Meadows	93.45	89.40	97.81	97.66
Gravel	1.14	17.63	78.35	92.73
Trees	1.60	70.46	90.34	83.67
Metal sheets	0.89	97.92	97.55	99.22
Bare soil	2.07	16.01	74.47	96.52
Bitumen	1.50	15.64	96.20	97.59
Bricks	0.92	67.76	90.96	93.14
Shadows	1.80	86.17	96.88	74.76
AA	11.75	59.51	90.37	92.53
OA	41.71	69.63	91.82	95.41
Kappa	0.87	58.48	89.06	93.89
SLIC Time	-	-	8.30	8.30
Sensing Time	3.08	3.08	3.15	3.15
Reconstruction Time	2370.08	2644.08	3201.77	3.02
Classification Time	88.74	16.51	14.01	1.83
Total Time	2461.89	2663.66	3227.22	16.30

Figure 23: Visual classification results on Pavia Center dataset

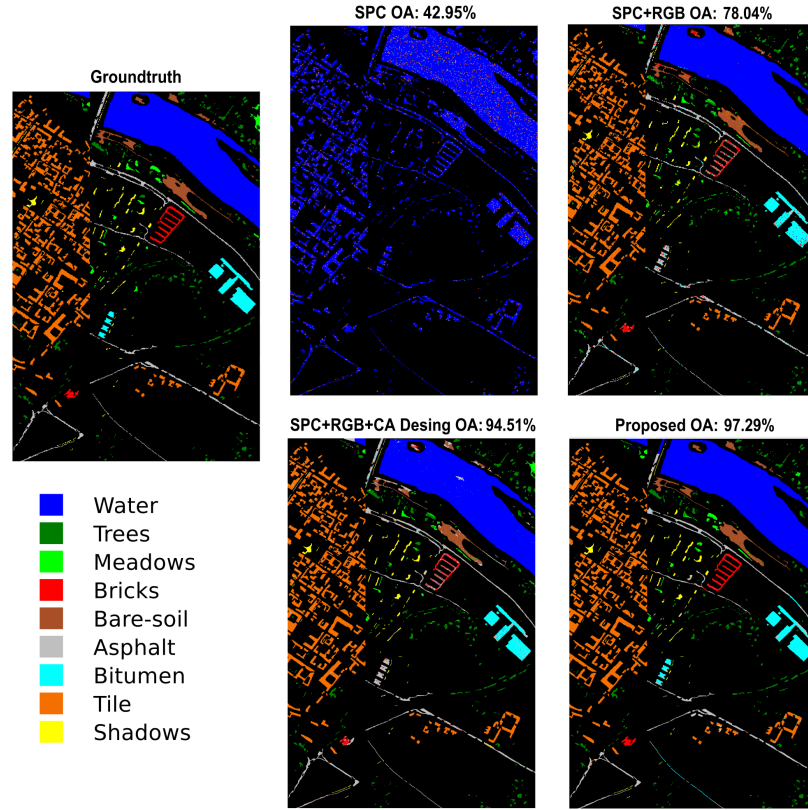


Table 5: Quantitative results on Pavia Center dataset (OA %)

Class	SPC	SPC+RGB	SPC+RGB+CA Design	Proposed
Water	91.36	99.79	99.34	100.00
Trees	0.74	87.85	95.62	90.94
Meadows	1.36	48.70	54.76	76.46
Bricks	0.90	58.52	60.04	97.50
Bare soil	0.43	80.21	73.93	96.33
Asphalt	1.42	85.39	92.48	90.09
Bitumen	0.87	79.22	77.27	92.74
Tile	7.18	95.57	98.73	98.88
Shadows	0.14	80.85	88.30	86.62
AA	11.60	79.57	82.28	92.17
OA	42.95	93.04	94.51	97.29
Kappa	0.98	90.16	92.23	96.16
SLIC Time	-	-	35.10	35.10
Sensing Time	2.62	2.62	11.02	11.02
Reconstruction Time	1950.15	2281.04	9895.51	11.20
Classification Time	74.41	9.08	59.41	7.41
Total Time	2027.18	2292.74	10001.04	64.74

Figure 24: Visual classification results on Salinas Valley dataset

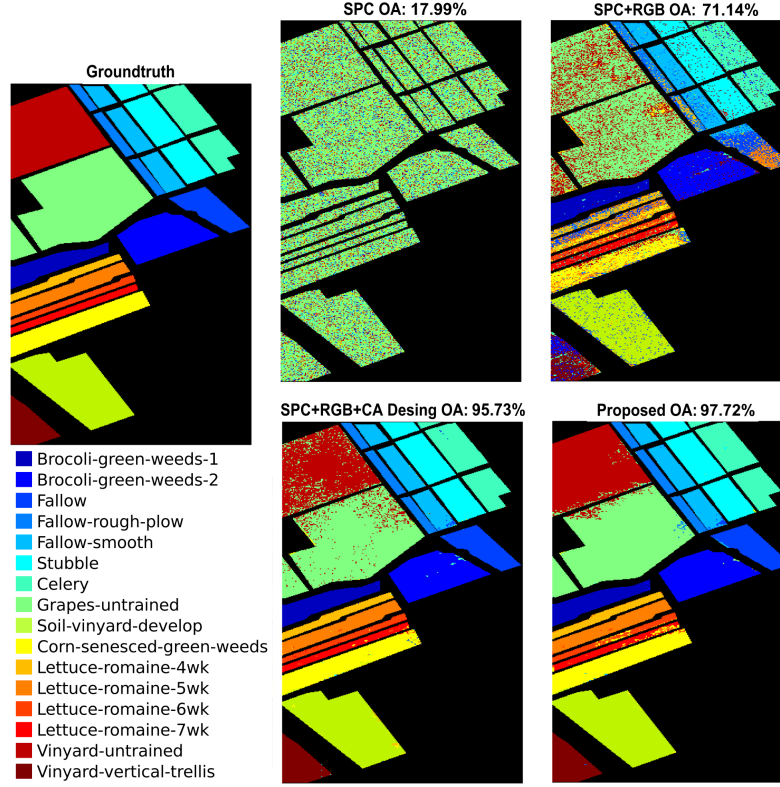


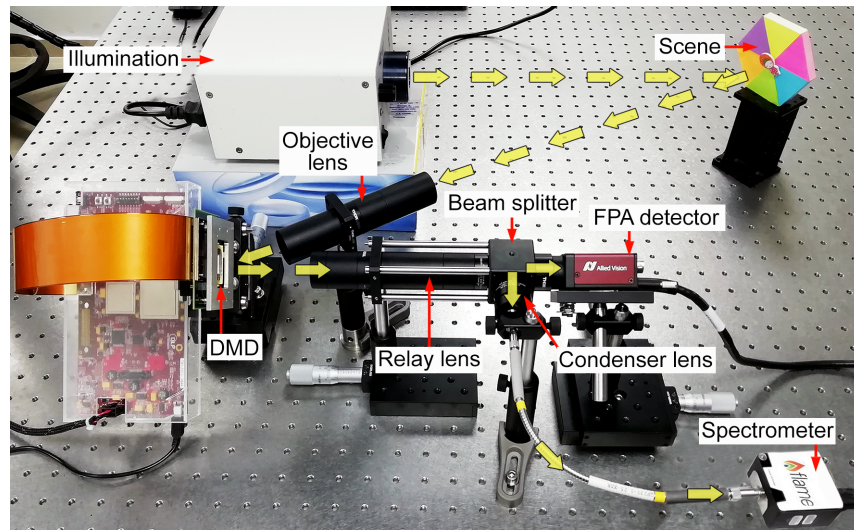
Table 6: Quantitative results on Salinas Valley dataset (OA %)

Class	Random	SPC+RGB	SPC+RGB+CA Design	Proposed
Broccoli-green-1	2.79	92.68	99.85	99.30
Broccoli-green-2	3.17	89.29	99.09	99.60
Fallow	2.83	48.79	<u>99.85</u>	98.94
Fallow-rough-plow	2.44	60.55	<u>98.78</u>	99.07
Fallow-smooth	4.11	87.27	97.31	<u>96.60</u>
Stubble	4.45	92.47	99.87	<u>99.37</u>
Celery	3.13	90.44	99.64	<u>99.50</u>
Grapes-untrained	67.03	73.13	<u>92.40</u>	97.03
Soil-vinyard-develop	8.25	89.25	<u>98.55</u>	99.56
Corn-senesced-green-weeds	2.14	62.11	<u>95.79</u>	97.86
Lettuce-romaine-4wk	2.43	59.18	98.88	<u>95.97</u>
Lettuce-romaine-5wk	2.85	48.26	<u>98.75</u>	98.91
Lettuce-romaine-6wk	3.28	73.47	94.21	86.24
Lettuce-romaine-7wk	2.43	78.88	94.39	<u>81.12</u>
Vinyard-untrained	10.35	30.56	<u>87.23</u>	97.08
Vinyard-vertical-trellis	2.66	64.14	<u>99.78</u>	99.89
AA	7.77	71.28	97.15	96.63
OA	17.99	71.14	<u>95.73</u>	97.72
Kappa	1.84	67.81	<u>95.24</u>	97.46
SLIC Time	-	-	4.53	4.53
Sensing Time	<u>3.31</u>	<u>3.31</u>	3.27	3.27
Reconstruction Time	<u>2398.71</u>	2686.37	2572.74	3.49
Classification Time	<u>153.32</u>	<u>37.00</u>	37.11	2.39
Total Time	<u>2555.34</u>	2726.67	2617.66	13.69

4.2. EXPERIMENTAL SETUP

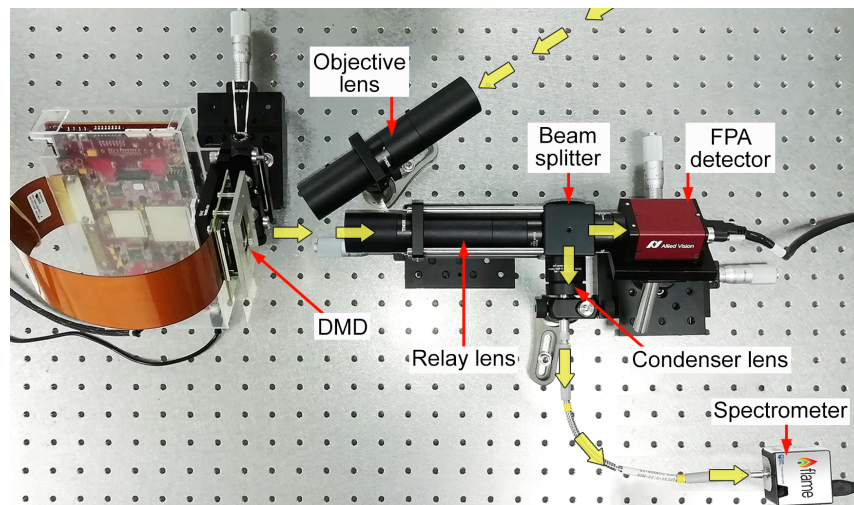
An optical testbed implementation of the proposed architecture was constructed in the laboratory to experimentally verify its performance. This prototype is shown in Fig 25. It is composed of two 100 mm objective lenses, a non-polarizing Beam-splitter to split the light, a digital micromirror device (DMD) as encoding element, a 100 mm relay lens, a F220SMA-A condenser connected through a fiber to an Ocean Optics Flame S-VIS-NIR-ES spectrometer, which is used as single pixel detector, and an AVT STINGRAY F-080B camera to acquire the RGB image, that is, at the wavelengths 455 nm, 540 nm and 656 nm. In Fig. 26 a top view of the implemented testbed is shown.

Figure 25: Testbed implemented to acquire the compressive spectral measurements SPC and side information from RGB.



In Fig. 26 a top view of the two arms that compose the acquisition architecture implemented is observed.

Figure 26: Top view of the implemented architecture.



4.3. ACQUIRED DATASETS

The cocoa bean is the dried and fermented seed of the *Theobroma cacao* tree (Fig 27), the raw material of the chocolate industry, and the main livelihood of more than 40 million people around the world ⁵⁴. Growing cocoa requires hot and rainy weather, which is common in tropical countries ⁵⁵. Africa leads the production of cocoa seed with 68% of the world's bean. Latin American countries have a 17% of production, led by Brazil, Ecuador, Mexico, Peru, the Dominican Republic, and Colombia, respectively. Cocoa ranks fifth among the most representative agricultural products of the Colombian economy. Santander is the number 1 department in cocoa production in the country with 46% of the total. Colombia needs to improve the techniques of certified cocoa classification to venture into emerging international markets ⁵⁶. Although the production of this bean has significantly increased from 3.3 megatons [Mt] in 2000 to 4.5 Mt in 2016 in the world, leaders of the chocolate industry project a global deficit of cocoa beans by 2020.

Figure 27: Process of the cob to the fermented cocoa beans.



This deficit represents an opportunity to increase exports. One of the strategies suggested by the International Cocoa Organization (ICCO) for this purpose is to innovate in tools that guarantee the classification of beans in a less complex and non-subjective way.

The Colombian technical standard NTC 1252 ⁵⁷ describes in detail the twelve chemical, physical, and sensory stages with which the national cocoa federation of Colombia (Fedecacao) defines the quality of dry cocoa beans today. Specifically, this process requires tasters, technical personnel and specialized equipment, and involves the destruction of the samples. In Fig 28 some stages of the traditional process of classification of cocoa beans in Colombia are shown, which is similar to other countries in the world.

Taking into account the global trends in food quality control, from the acquisition and processing of spatial spectral information of them, this can be an optimal alternative for the classification of cocoa beans non-invasively, without waste, accurate and faster.

For this reason, dry cocoa beans are the material chosen to test the performance of the classification system proposed in this work. It should be noted that this proposal is approved by the national federation of cacao growers of Colombia (Fedecacao) and was

⁵⁴ Rolando Saltini, Renzo Akkerman, and Stina Frosch. "Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality". In: *Food control* 29.1 (2013), pp. 167–187.

⁵⁵ Mohd Shavez Beg et al. "Status, supply chain and processing of cocoa-A review". In: *Trends in food science & technology* 66 (2017), pp. 108–116.

⁵⁶ John Edem Kongor et al. "Constraints for future cocoa production in Ghana". In: *Agroforestry systems* (2017), pp. 1–13.

⁵⁷ Norma Técnica Colombiana. *NTC 1252 clasificación del cacao en grano destinado a la industrialización para consumo humano*. 2003.

Figure 28: Some stages of the traditional classification of cocoa bean in Colombia.



Source: *Mercadeo Internacional del cacao en grano y las potencialidades en Colombia*, 2016

selected by the Government of Santander and Colciencias to grant a National Master's Scholarship that covered the expenses of the author during his postgraduate studies. In addition, this idea was one of the two winners in Colombia of the Bayer Young Community Innovators 2017 competition. This championship is promoted by the multinational company Bayer S.A., and awards young people with innovative initiatives that are replicable, of high impact for agriculture or health, and faithful to the company slogan Science for a Better Life⁵⁸⁵⁹.

The cocoa beans used for this work were grown and fermented by a local producer in Rionegro, Santander, Colombia (7° 15' 51" N, 73° 08' 58" W). Each bean was classified by Fedecacao using the traditional method in NTC 1252 depending on the degree of fermentation. Expressly, three categories of beans were selected: low fermentation, correct fermentation, high fermentation. The seeds were taken to the optical laboratory of HDSP at the UIS, Bucaramanga, on the same day that Fedecacao classified them.

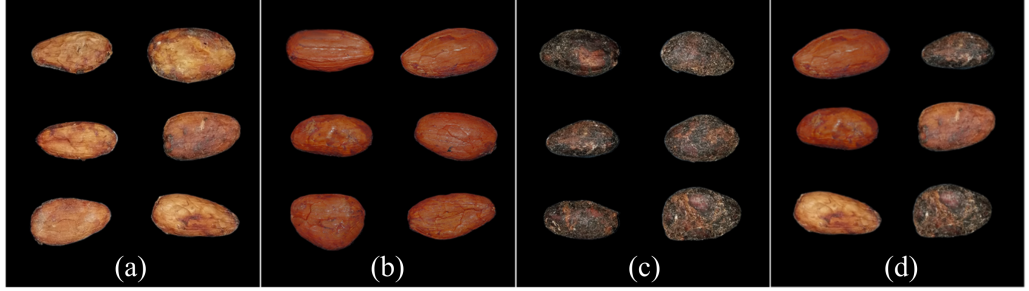
For this research work were supplied 108 beans, 36 of each of the three fermentation categories. Five black cards were arranged for each fermentation category, each with 6 grains of the same class. Additionally, three black cards each with 6 mixed grains (2 of each fermentation class per card) were fixed as shown in Fig. 29. In total, 18 scenes of 6 beans each were arranged for acquisition with the optical architecture described in section 4.2. Measurements of the target scenes were captured with this system for a compression ratio $\gamma = 0.5$, and $M = N = 384$.

As one of the results of this research proposal, 36 spatio-spectral datasets with 108 cocoa beans were acquired. Specifically, a total of 18 scenes were captured with measures compressed by single-pixel optical architecture, and 18 RGB images of the same scenes. The datasets (in SPC and RGB) were acquired under 23°C of temperature and 60% humidity conditions. Note that, as shown in the implemented architecture (Fig. 25), the coded aperture (DMD) is located before the light bifurcation with the beam splitter, which guarantees the registration of the SPC and RGB images. In this way, the datasets are suitable for fusion tests without the need to include a stage of registration of both datasets (SPC and RGB).

⁵⁸ Bayer Andina. *Los estudiantes más innovadores de Colombia*. url <https://bayerinnovacion.wordpress.com/>. 2017.

⁵⁹ Bayer S.A. *Innovadores al servicio de la comunidad*. url <https://andina.bayer.com/es/sala-de-prensa/dos-nuevos-lideres-sociales-en-colombia-seran-apoyados-por-bayer.php>. 2017.

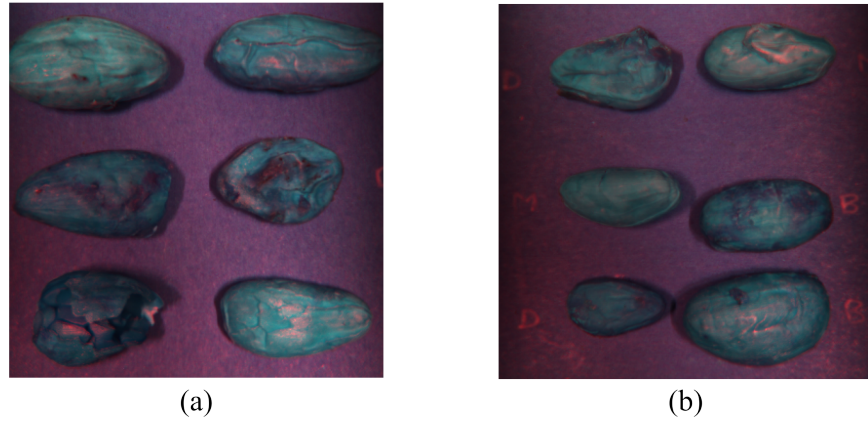
Figure 29: Cocoa beans of (a) low fermentation, (b) correct fermentation, (c) high fermentation, and (d) mixed beans.



4.4. ANALYSIS OF EXPERIMENTAL RESULTS

The experiments and analysis performed in section 4.1 with the Pavia University, Pavia Center and Salinas images, were also developed with some of the datasets acquired in the laboratory through the implemented architecture. Specifically, the spatio-spectral information of two mixed beans cards (as in Fig. 4.11 (d)) were used to test the development of the implemented algorithm. Figure 30 shows the two mixed cards of beans selected to realize the simulations and to evaluate the classification performance. Each of these databases has a spatial resolution of 384×384 pixels.

Figure 30: False color composite of mixed cocoa beans (a) dataset 1, and (b) dataset 2.

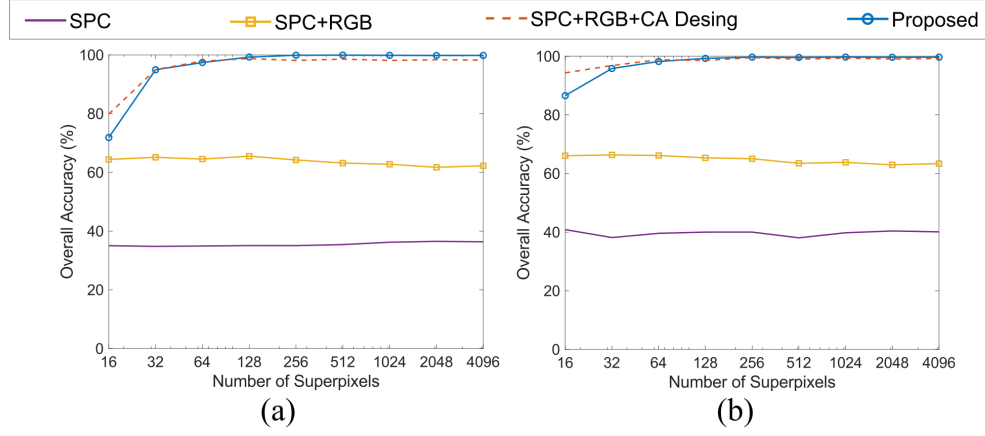


4.4.1. Analysis of superpixels number and compression rates. In this experiment, the same superpixel numbers were varied as an input parameter to evaluate the accuracy of the classification. Figure 31 shows the results obtained with each approach of Fig. 20.

Each tested scene includes 2 cocoa beans per category: low fermentation, correct fermentation, and high fermentation. Similar results are obtained with the two datasets. The proposed method improves its performance exponentially when N_{seg} increases. Note that the number of pixels in a RGB band is $384 \times 384 = 147,456$ pixels. And that, when the SLIC algorithm associates all these pixels in 128 groups, the clustering is sufficient for the accuracy of the classification to approach 100 %. According to this, for application purposes, 256 superpixels would guarantee high classification accuracy.

On the other hand, the approach that reconstructs SPC + RGB + CA design achieves

Figure 31: Overall accuracy of the classification by varying the number of superpixels (N_{seg}) in (a) Cocoa bean 1 dataset, and (b) Cocoa bean 2 dataset.



results very similar to those of the proposed method. In contrast, the SPC, and SPC + RGB methods show low OAs of 39% and 62%, respectively.

4.4.2. Visual maps and Quantitative results. After analyzing the behavior of the classification with each approach by varying the Nseg parameter, it is convenient to graph the maps of class assignment. To do this, $N_{seg} = 512$ was chosen for comparison purposes, since in this value the results with all the approaches have stabilized. Figures (4.13) - (4.14) and Tables (4.4)-(4.5) shows the visual and quantitative results.

In this case, to evaluate the accuracy of the SVM classifier on the measurements extracted or reconstructed with each method, a ground truth for each dataset had to be constructed from the information provided by Fedecacao. It should be remembered that the selection process was carried out by specialized technical personnel of the federation, who assigned a class to each grain, before being taken to the optical laboratory of HDSP, UIS.

In Fig 32 and Table 4.4 of the mixed board 1, it is notable that the cocoa beans with High fermentation were the best identified (100%), followed by the correctly fermented beans with a OA similar value (99.92%), and finally, but also with satisfactory results, the low-fermented beans (99.85%). The results of the classification with the mixed board 2 are similar to the board 1. A high precision is obtained with the proposed method (Fig 33 and Table 4.5). In both acquired datasets it is remarkable that the SPC + RGB + CA design, provides a high classification performance. However, the total time with the proposed approach is significantly less, representing a saving of around 99 minutes compared with the third approach.

Figure 32: Visual classification results on Cocoa bean 1 dataset

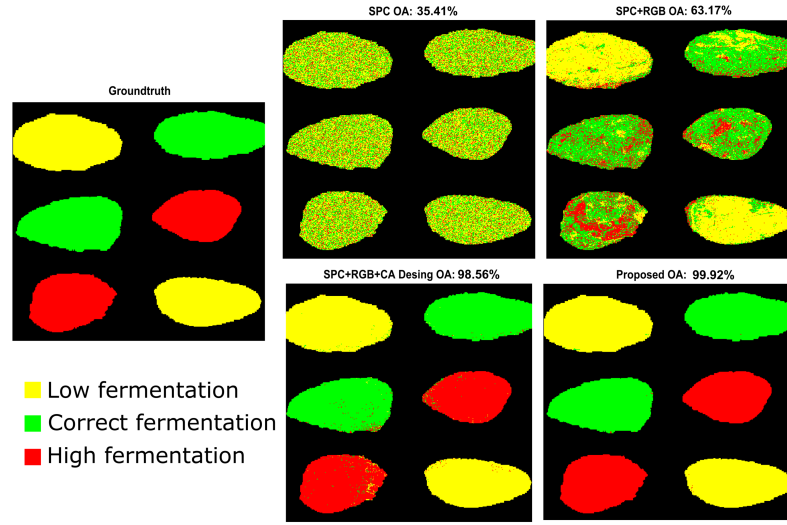


Table 7: Quantitative results on Cacao bean 1 dataset (OA %)

Class	Random	SPC+RGB	SPC+RGB+CA Design	Proposed
Low Fermentation	49.25	83.64	<u>99.40</u>	99.85
Correct Fermentation	36.64	66.75	<u>98.64</u>	99.92
High Fermentation	16.93	33.70	<u>97.42</u>	100.00
AA	34.27	61.37	<u>98.49</u>	99.92
OA	35.41	63.17	<u>98.56</u>	99.92
Kappa	1.47	44.01	<u>97.83</u>	99.88
SLIC Time	-	-	8.55	8.55
Sensing Time	<u>8.07</u>	<u>8.07</u>	3.97	3.97
Reconstruction Time	<u>5740.02</u>	8512.68	6632.35	7.44
Classification Time	153.55	104.84	<u>42.79</u>	2.15
Total Time	<u>5901.64</u>	8625.60	6687.67	22.11

Figure 33: Visual classification results on the Cocoa bean 2 dataset

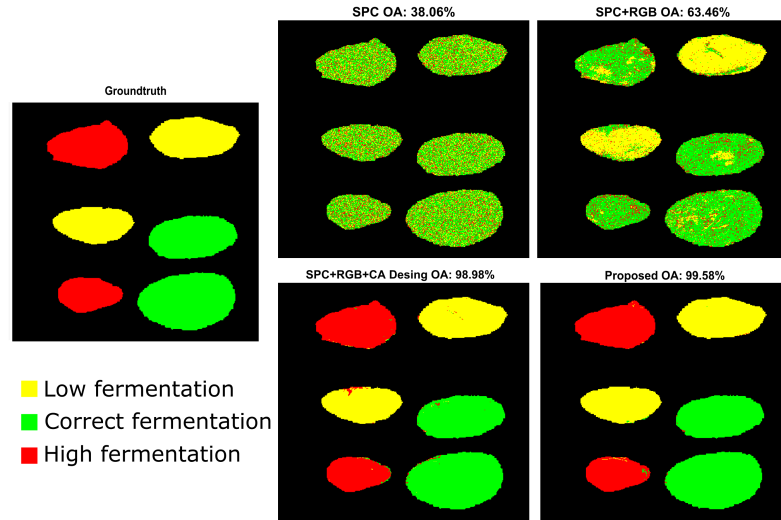


Table 8: Quantitative results on Cacao bean 2 dataset (OA %)

Class	Random	SPC+RGB	SPC+RGB+CA Design	Proposed
Low Fermentation	25.19	83.68	<u>98.81</u>	99.64
Correct Fermentation	57.50	77.07	<u>99.22</u>	99.83
High Fermentation	20.33	17.74	<u>98.80</u>	99.12
AA	34.34	59.50	<u>98.94</u>	99.53
OA	38.06	63.46	<u>98.98</u>	99.58
Kappa	1.70	41.79	<u>98.44</u>	99.36
SLIC Time	-	-	6.44	6.44
Sensing Time	7.14	7.14	<u>7.32</u>	<u>7.32</u>
Reconstruction Time	<u>5125.68</u>	5898.77	5943.89	7.64
Classification Time	<u>77.41</u>	34.00	<u>9.74</u>	1.46
Total Time	<u>5210.24</u>	5939.91	5967.38	22.86

5. CONCLUSIONS

A new supervised classification method for spectral images acquired with SPC and RGB, which performs the classification task directly on the compressed measurements, was developed in this research work. The Single Pixel Camera acquisition model was used since it allows to encode the spatial bands before performing as spectral-wise integration. In the proposed method, a pattern to add the side information based on superpixels from an RGB in SPC was first designed. After stacking the RGB in SPC measurements, a mathematical model is developed to reconstruct the fusion.

The main contribution of this work is the coding pattern design based on Hadamard matrices. The proposed design takes into account the spatial information of an RGB side image, through superpixels, to design the coded apertures to sense with SPC. On the other hand, the proposed design realized the task of classification of fusion compressed measurements without includes a reconstruction step. For this, the proposed approach classifies the spectral signatures of the superpixels.

The supervised classification for compressed measurements was validated trough several experiments. One hyperspectral remote sensing scene from the AVIRIS sensor and two from the ROSIS-03 sensor were used in the experiments. In addition, an architecture was implemented and spatial-spectral information of 108 cocoa beans was acquired. The proposed algorithm was validated with the images of two cocoa beans scenes each with six grains. In general, the results show that the classification directly with the extracted features provides similar accuracy results in comparison with those provided by executing the classification of the reconstruction of SPC + RGB when the coded aperture proposed in this work was used. Particularly, a difference of at most 36% in terms of overall accuracy was observed when comparing the classification results obtained by SPC acquired with randomly coding pattern + RGB. In terms of the total time of execution, the mathematical approach proposed in this book presents significant and promising results that will continue to be studied by the author.

BIBLIOGRAPHY

Abdul, J. "The Shannon Sampling Theorem—Its Various Extensions and Applications: A". In: *Proceedings of the IEEE* 65.11 (1977), p. 1565 (cit. on p. 17).

Achanta, Radhakrishna et al. "SLIC superpixels compared to state-of-the-art superpixel methods". In: *IEEE transactions on pattern analysis and machine intelligence* 34.11 (2012), pp. 2274–2282 (cit. on pp. 27, 28, 41).

Andina, Bayer. *Los estudiantes más innovadores de Colombia*. url <https://bayerinnovacion.wordpress.com>. 2017 (cit. on p. 54).

Arce, Gonzalo R et al. "Compressive coded aperture spectral imaging: An introduction". In: *Signal Processing Magazine, IEEE* 31.1 (2014), pp. 105–115 (cit. on pp. 13, 19–21).

Arguello, Henry and Gonzalo R Arce. "Colored coded aperture design by concentration of measure in compressive spectral imaging". In: *IEEE Transactions on Image Processing* 23.4 (2014), pp. 1896–1908 (cit. on p. 13).

Arguello, Henry, Claudia V Correa, and Gonzalo R Arce. "Fast lapped block reconstructions in compressive spectral imaging". In: *Applied Optics* 52.10 (2013), pp. D32–D45 (cit. on p. 42).

Baldeck, Claire A and Gregory P Asner. "Estimating vegetation beta diversity from airborne imaging spectroscopy and unsupervised clustering". In: *Remote Sensing* 5.5 (2013), pp. 2057–2071 (cit. on p. 18).

Beg, Mohd Shavez et al. "Status, supply chain and processing of cocoa-A review". In: *Trends in food science & technology* 66 (2017), pp. 108–116 (cit. on p. 53).

Bioucas-Dias, José M et al. "Hyperspectral remote sensing data analysis and future challenges". In: *IEEE Geoscience and remote sensing magazine* 1.2 (2013), pp. 6–36 (cit. on p. 12).

Bovolo, Francesca et al. *A novel context-sensitive SVM for classification of remote sensing images*. University of Trento, 2006 (cit. on p. 19).

-
- Bruzzone, Lorenzo, Mingmin Chi, and Mattia Marconcini. "A novel transductive SVM for semisupervised classification of remote-sensing images". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.11 (2006), pp. 3363–3373 (cit. on p. 19).
- Camps-Valls, Gustavo et al. "Composite kernels for hyperspectral image classification". In: *IEEE geoscience and remote sensing letters* 3.1 (2006), pp. 93–97 (cit. on p. 19).
- Cao, Xun et al. "A prism-mask system for multispectral video acquisition". In: *IEEE transactions on pattern analysis and machine intelligence* 33.12 (2011), pp. 2423–2435 (cit. on p. 21).
- Chang, Chein-I. *Hyperspectral data exploitation: theory and applications*. John Wiley & Sons, 2007 (cit. on p. 16).
- *Hyperspectral imaging: techniques for spectral detection and classification*. Vol. 1. Springer Science & Business Media, 2003 (cit. on p. 13).
- Colombiana, Norma Técnica. *NTC 1252 clasificación del cacao en grano destinado a la industrialización para consumo humano*. 2003 (cit. on p. 53).
- Donoho, David L and Xiaoming Huo. "Uncertainty principles and ideal atomic decomposition". In: *Information Theory, IEEE Transactions on* 47.7 (2001), pp. 2845–2862 (cit. on p. 20).
- Donoho, David L et al. "Compressed sensing". In: *IEEE Transactions on information theory* 52.4 (2006), pp. 1289–1306 (cit. on pp. 13, 19).
- Dougherty, James, Ron Kohavi, and Mehran Sahami. "Supervised and unsupervised discretization of continuous features". In: *Machine Learning Proceedings 1995*. Elsevier, 1995, pp. 194–202 (cit. on p. 13).
- Duarte, Marco F et al. "Single-pixel imaging via compressive sampling". In: *IEEE signal processing magazine* 25.2 (2008), pp. 83–91 (cit. on pp. 14, 21).
- Eismann, Michael T and Russell C Hardie. "Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.3 (2005), pp. 455–465 (cit. on p. 32).
- Fauvel, Mathieu et al. "Advances in spectral-spatial classification of hyperspectral images". In: *Proceedings of the IEEE* 101.3 (2012), pp. 652–675 (cit. on p. 18).

-
- Fauvel, Mathieu et al. "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles". In: *IEEE Transactions on Geoscience and Remote Sensing* 46.11 (2008), pp. 3804–3814 (cit. on p. 19).
- Garcia, Hans et al. "Multi-Resolution Coded Apertures Based on Side Information for Single Pixel Spectral Reconstruction". In: *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE. 2018, pp. 2215–2219 (cit. on pp. 14, 24).
- Garcia, Hans et al. "Multi-resolution reconstruction algorithm for compressive single pixel spectral imaging". In: *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE. 2017, pp. 468–472 (cit. on p. 41).
- Ghamisi, Pedram et al. "Advanced spectral classifiers for hyperspectral images: A review". In: *IEEE Geoscience and Remote Sensing Magazine* 5.1 (2017), pp. 8–32 (cit. on p. 18).
- Gu, Bin and Victor S Sheng. "A Robust Regularization Path Algorithm for ν -Support Vector Classification". In: *IEEE Transactions on neural networks and learning systems* 28.5 (2017), pp. 1241–1248 (cit. on p. 29).
- Gualtieri, J Anthony and Robert F Crompt. "Support vector machines for hyperspectral remote sensing classification". In: *The 27th AIPR Workshop: Advances in Computer-Assisted Recognition*. International Society for Optics and Photonics. 1999, pp. 221–232 (cit. on p. 18).
- Hagen, Nathan A and Michael W Kudenov. "Review of snapshot spectral imaging technologies". In: *Optical Engineering* 52.9 (2013), p. 090901 (cit. on pp. 12, 17, 19).
- He, Zhi et al. "Low-rank group inspired dictionary learning for hyperspectral image classification". In: *Signal Processing* 120 (2016), pp. 209–221 (cit. on p. 28).
- Heald, Mark A and Jerry B Marion. *Classical electromagnetic radiation*. Courier Corporation, 2012 (cit. on p. 16).
- Hinojosa, Carlos A et al. *Compressive spectral imaging using multiple snapshot colored-mosaic detector measurements*. 2016. DOI: [10.1117/12.2224369](https://doi.org/10.1117/12.2224369) (cit. on p. 21).
- James, Alex Pappachen and Belur V Dasarathy. "Medical image fusion: A survey of the state of the art". In: *Information Fusion* 19 (2014), pp. 4–19 (cit. on p. 12).

-
- Keller, James M, Michael R Gray, and James A Givens. "A fuzzy k-nearest neighbor algorithm". In: *IEEE transactions on systems, man, and cybernetics* 4 (1985), pp. 580–585 (cit. on p. 18).
- Kongor, John Edem et al. "Constraints for future cocoa production in Ghana". In: *Agroforestry systems* (2017), pp. 1–13 (cit. on p. 53).
- Landau, HJ. "Sampling, data transmission, and the Nyquist rate". In: *Proceedings of the IEEE* 55.10 (1967), pp. 1701–1706 (cit. on p. 17).
- Li, Jun et al. "Generalized composite kernel framework for hyperspectral image classification". In: *IEEE transactions on geoscience and remote sensing* 51.9 (2013), pp. 4816–4829 (cit. on p. 29).
- Lin, Fu, Mihailo R Jovanovic, and Tryphon T Georgiou. "An ADMM algorithm for matrix completion of partially known state covariances". In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE. 2013, pp. 1684–1689 (cit. on p. 66).
- Lin, Xing et al. "Dual-coded compressive hyperspectral imaging". In: *Optics letters* 39.7 (2014), pp. 2044–2047 (cit. on pp. 14, 21).
- Lin, Xing et al. "Spatial-spectral encoded compressive hyperspectral imaging". In: *ACM Transactions on Graphics (TOG)* 33.6 (2014), p. 233 (cit. on p. 21).
- Loncan, Laetitia et al. "Hyperspectral pansharpener: A review". In: *arXiv preprint arXiv:1504.04531* (2015) (cit. on pp. 24, 32).
- Lorente, D et al. "Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment". In: *Food and Bioprocess Technology* 5.4 (2012), pp. 1121–1142 (cit. on p. 12).
- Martín, Gabriel and José M Bioucas-Dias. "Hyperspectral blind reconstruction from random spectral projections". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9.6 (2016), pp. 2390–2399 (cit. on pp. 12, 18).
- Mountrakis, Giorgos, Jungho Im, and Caesar Ogole. "Support vector machines in remote sensing: A review". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 66.3 (2011), pp. 247–259 (cit. on p. 29).

-
- Nasrabadi, Nasser M. "Hyperspectral target detection: An overview of current and future challenges". In: *IEEE Signal Processing Magazine* 31.1 (2013), pp. 34–44 (cit. on p. 12).
- Parikh, Neal, Stephen Boyd, et al. "Proximal algorithms". In: *Foundations and Trends in Optimization* 1.3 (2014), pp. 127–239 (cit. on p. 66).
- Ramirez, Ana et al. "Spectral image classification from optimal coded-aperture compressive measurements". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6 (2014), pp. 3299–3309 (cit. on p. 22).
- Rueda-Chacón, Hoover Fabián, Cesar Augusto Vargas-García, and Henry Arguello-Fuentes. "Single-pixel optical sensing architecture for compressive hyperspectral imaging". In: *Revista Facultad de Ingeniería Universidad de Antioquia* 73 (2014), pp. 124–133 (cit. on pp. 14, 21).
- S.A., Bayer. *Innovadores al servicio de la comunidad*. url <https://andina.bayer.com/es/sala-de-prensa/dos-nuevos-lideres-sociales-en-colombia-seran-apoyados-por-bayer.php>. 2017 (cit. on p. 54).
- Saltini, Rolando, Renzo Akkerman, and Stina Frosch. "Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality". In: *Food control* 29.1 (2013), pp. 167–187 (cit. on p. 53).
- Sanchez, Karen, Carlos Hinojosa, and Henry Arguello. "Supervised Classification of Hyperspectral Images using Side Information". In: *Laser Applications to Chemical, Security and Environmental Analysis*. Optical Society of America. 2018, JW5E–5 (cit. on p. 26).
- "Supervised spatio-spectral classification of fused images using superpixels". In: *Applied optics* 58.7 (2019), B9–B18 (cit. on p. 26).
- Scarlett, Jonathan, Jamie S Evans, and Subhrakanti Dey. "Compressed sensing with prior information: Information-theoretic limits and practical decoders". In: *IEEE Transactions on Signal Processing* 61.2 (2012), pp. 427–439 (cit. on pp. 14, 24).
- Scholkopf, Bernhard, Alexander J Smola, and Francis Bach. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. The MIT Press, 2018 (cit. on p. 29).

-
- Shaw, Gary A and Hsiao-hua K Burke. "Spectral imaging for remote sensing". In: *Lincoln Laboratory Journal* 14.1 (2003), pp. 3–28 (cit. on p. 13).
- Thanh Noi, Phan and Martin Kappas. "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery". In: *Sensors* 18.1 (2018), p. 18 (cit. on p. 18).
- Vargas, Edwin et al. "Spectral image fusion from compressive measurements". In: *IEEE Transactions on Image Processing* 28.5 (2019), pp. 2271–2282 (cit. on pp. 14, 25).
- Wald, George. "Human vision and the spectrum." In: *Science* (1945) (cit. on p. 16).
- Wei, Q. et al. "Multiband Image Fusion Based on Spectral Unmixing". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.12 (2016), pp. 7236–7249. DOI: [10.1109/TGRS.2016.2598784](https://doi.org/10.1109/TGRS.2016.2598784) (cit. on p. 24).
- Wei, Qi. "Bayesian fusion of multi-band images: a powerful tool for super-resolution". PhD thesis. Institut national polytechnique de Toulouse (INPT), 2015 (cit. on pp. 24, 25).
- Yokoya, Naoto, Claas Grohnfeldt, and Jocelyn Chanussot. "Hyperspectral and multispectral data fusion: A comparative review of the recent literature". In: *IEEE Geoscience and Remote Sensing Magazine* 5.2 (2017), pp. 29–56 (cit. on p. 13).
- Yuan, Xin, Yangyang Sun, and Shuo Pang. "Compressive video sensing with side information". In: *Applied optics* 56.10 (2017), pp. 2697–2704 (cit. on p. 24).
- Zhang, Xuewen et al. "SLIC superpixels for efficient graph-based dimensionality reduction of hyperspectral imagery". In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI*. Vol. 9472. International Society for Optics and Photonics. 2015, p. 947209 (cit. on p. 27).

APPENDIX

ADMM ALGORITHM TO SOLVE THE PROPOSED OPTIMIZATION PROBLEM

The minimization in Eq. (2.10) can be efficiently solved by the well-known alternating direction method of multipliers (ADMM). First, an auxiliary matrix $\mathbf{Z}$ with the same size of $\bar{\mathbf{F}}$ is used to separate the problem. In this way, problem in Eq. (2.10) becomes

$$\begin{aligned} \min_{\bar{\mathbf{F}}, \mathbf{Z}} \quad & \frac{1}{2} \|\mathbf{F}_h - \mathbf{D}_h \mathbf{U}^T \bar{\mathbf{F}}\|_F^2 + \lambda \|\mathbf{Z}\|_* \\ \text{s.t.} \quad & \mathbf{Z} = \bar{\mathbf{F}}. \end{aligned} \quad (1)$$

The augmented Lagrangian associated with the optimization in Eq. (1) can be written as

$$\mathcal{L}(\bar{\mathbf{F}}, \mathbf{Z}, \mathbf{G}) = \frac{1}{2} \|\mathbf{F}_h - \mathbf{D}_h \mathbf{U}^T \bar{\mathbf{F}}\|_F^2 + \lambda \|\mathbf{Z}\|_* + \frac{\rho}{2} \|\bar{\mathbf{F}} - \mathbf{Z} + \mathbf{G}\|_F^2, \quad (2)$$

where $\mathbf{G}$ is the scaled dual variable and $\rho > 0$ is the weighting of the augmented Lagrangian term Neal Parikh, Stephen Boyd, et al. “Proximal algorithms”. In: *Foundations and Trends in Optimization* 1.3 (2014), pp. 127–239. The ADMM solution of Eq. (1) is summarized in Algorithm 2 that consists of minimizing $\bar{\mathbf{F}}$, $\mathbf{Z}$, and $\mathbf{G}$, alternatively. Specifically, the minimization of the three matrices consists of the following updates:

- 1) *Updating $\bar{\mathbf{F}}$* : $\bar{\mathbf{F}}^{(t)}$ is obtained by minimizing $\mathcal{L}$ with respect to $\bar{\mathbf{F}}$, while $(\mathbf{Z}, \mathbf{G})$ are fixed. We calculate the derivative of $\mathcal{L}$ with respect to $\bar{\mathbf{F}}$ and set it to zero to obtain $\bar{\mathbf{F}}$ as follows:

$$\bar{\mathbf{F}}^{(t)} = \frac{\mathbf{U}^T \mathbf{D}_h^T \mathbf{F}_h + \rho \mathbf{I}(\mathbf{Z} - \mathbf{G})}{\rho \mathbf{I} + \mathbf{U}^T \mathbf{D}_h^T \mathbf{D}_h \mathbf{U}} \quad (3)$$

- 2) *Updating $\mathbf{Z}$* : the update of $\mathbf{Z}$ is given by the singular value thresholding operator Fu Lin, Mihailo R Jovanovic, and Tryphon T Georgiou. “An ADMM algorithm for matrix completion of partially known state covariances”. In: *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*. IEEE. 2013, pp. 1684–1689. Specifically, we compute the SVD of the symmetric matrix

$$\bar{\mathbf{F}} + \mathbf{G} = \hat{\mathbf{U}} \mathbf{\Sigma} \hat{\mathbf{U}}^* \quad (4)$$

and apply the soft-thresholding operator to the resulting singular values

$$S_{1/\rho}(\mathbf{\Sigma}) = \text{diag} \{ (\sigma_i - 1/\rho)_+ \}, \quad (5)$$

where $(a)_+ := \max \{a, 0\}$. Then, the update of $\mathbf{Z}$ is defined as

$$\mathbf{Z}^{(t)} = \hat{\mathbf{U}} S_{1/\rho}(\mathbf{\Sigma}) \hat{\mathbf{U}}^*. \quad (6)$$

After obtaininig the estimation of $\bar{\mathbf{F}}$, the rich spatio-spectral image can be obtained using Eq. (2.9).

Algorithm 2 ADMM sub-iterations to estimate $\bar{\mathbf{F}}$

Input: $\mathbf{F}_h, \lambda, \rho, \mathbf{D}_h, \mathbf{U}, MAXITER$.

Initialization :

1: $\mathbf{Z}^{(0)} = \mathbf{0}, \mathbf{G}^0 = \mathbf{0}$

Main iterations

2: **for** $t = 1$ to $MAXITER$ **do**

3: $\bar{\mathbf{F}}^{(t)} \leftarrow \arg \min_{\bar{\mathbf{F}}} \mathcal{L}(\bar{\mathbf{F}}, \mathbf{Z}^{(t-1)}, \mathbf{G}^{(t-1)})$

4: $\mathbf{Z}^{(t)} \leftarrow \arg \min_{\mathbf{Z}} \mathcal{L}(\bar{\mathbf{F}}^{(t)}, \mathbf{Z}, \mathbf{G}^{(t-1)})$

5: $\mathbf{G}^{(t)} \leftarrow \mathbf{G}^{(t-1)} + \bar{\mathbf{F}} - \mathbf{Z}$

6: **end for**
