

Adaptive Multisensor Acquisition via Spatial Contextual Information for Compressive Spectral Image Classification

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY-SA 4.0

SUBMISSION DATE / POSTED DATE

15-07-2021 / 17-07-2021

CITATION

Diaz, Nelson; Marcos, Juan; Vera, Esteban; Arguello, Henry (2021): Adaptive Multisensor Acquisition via Spatial Contextual Information for Compressive Spectral Image Classification. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.14988804.v1>

DOI

[10.36227/techrxiv.14988804.v1](https://doi.org/10.36227/techrxiv.14988804.v1)

Adaptive Multisensor Acquisition via Spatial Contextual Information for Compressive Spectral Image Classification

Nelson Diaz, *Member, IEEE*, Juan Ramirez, *Member, IEEE*, Esteban Vera, *Member, IEEE*,
and Henry Arguello, *Senior, IEEE*

Abstract—Spectral image classification uses the huge amount of information provided by spectral images to identify objects in the scene of interest. In this sense, spectral images typically contain redundant information that is removed in later processing stages. To overcome this drawback, compressive spectral imaging (CSI) has emerged as an alternative acquisition approach that captures the relevant information using a reduced number of measurements. Various methods that classify spectral images from compressive projections have been recently reported whose measurements are captured by non-adaptive, or adaptive schemes discarding any contextual information that may help to reduce the number of captured projections. In this paper, an adaptive compressive acquisition method for spectral image classification is proposed. In particular, we adaptively design coded aperture patterns for a dual-arm CSI acquisition architecture, where the first system obtains compressive multispectral projections and the second arm registers compressive hyperspectral snapshots. The proposed approach exploits the spatial contextual information captured by the multispectral arm to design the coding patterns such that subsequent snapshots acquire the scene's complementary information improving the classification performance. Results of extensive simulations are shown for two state-of-the-art databases: Pavia University and Indian Pines. Furthermore, an experimental setup that performs the adaptive sensing was built to test the performance of the proposed approach on a real data set. The proposed approach exhibits superior performance with respect to other methods that classify spectral images from compressive measurements.

Index Terms—Compressive spectral imaging, adaptive acquisition, spatial contextual information, spectral image classification.

I. INTRODUCTION

SPECTRAL image (SI) classification is an important topic in remote sensing that aims at assigning predefined labels

Part of this work was supported by the Vicerrectoría de Investigación y Extensión at the Universidad Industrial de Santander through the project entitled: “Cámara difractiva compacta y de bajo costo para la adquisición comprimida de imágenes espectrales en la industria agrícola santanderena” under the grant 2699. The work of the N. Diaz was supported by the doctoral grant of Colciencias 727, doctorados nacionales de 2015.

N. Diaz is with the Department of Electrical Engineering, Universidad Industrial de Santander, Bucaramanga, Santander, 680002 Colombia, e-mail: nelson.diaz@saber.uis.edu.co.

J. Ramirez is with the Department of Computer Science, Universidad Rey Juan Carlos, Campus Móstoles, Spain, e-mail: juanmarcos.ramirez@urjc.es.

E. Vera is with the School of Electrical Engineering, Pontificia Universidad Católica de Valparaíso, Chile e-mail: esteban.vera@pucv.cl.

H. Arguello is with the Department of Computer Science, Universidad Industrial de Santander, Bucaramanga, Santander, 680002, Colombia, e-mail: henarfu@uis.edu.co.

Manuscript received April 19, 2005; revised August 26, 2015.

to the corresponding SI pixels. This approach has been widely used to detect and characterize materials in a scene of interest because of the huge amount of spatial-spectral information provided by SIs. In particular, SI classification has been used in various remote sensing applications such as forest classification [1], monitoring of agricultural land use [2], and precision agriculture [3]. In the last two decades, different pixel-based classification approaches have been proposed to label SIs, including methods based on support vector machines (SVMs) [4], neural networks [5], dictionary-based sparse representation [6], and multinomial logistic regression [7].

In general, pixel-based classification techniques suffer the so-called curse of dimensionality [8] that degrades the labeling performance as the number of spectral bands increases. To overcome this drawback, various SI classification approaches firstly resort to either dimensionality reduction (DR) or feature extraction techniques that obtain classification attributes based on the assumption that the relevant information embedded in spectral signatures lies in a lower-dimensional space. In this sense, principal component analysis (PCA) [9], independent component analysis (ICA) [10], and kernel-based methods [11], [12] are just few examples of DR approaches. Further, it is well-known that considering the spatial contextual information (SCI) improves the labeling accuracy, minimizing the salt and pepper noise in classification maps [7], [13], [14] [15]. Despite that, traditional pixel-based classifiers do not take contextual information into account.

On the other hand, spectral image classification methods are typically applied to data sets acquired by scanning sensors that obtain the entire image based on the Nyquist sampling theorem [16] leading to a high demand in storage and transmission needs during the acquisition process. To overcome these drawbacks, compressive spectral imaging (CSI) has emerged as a compressive sensing-based acquisition framework that captures the SI relevant information in a reduced set of random projections. The most representative CSI architecture was presented by Wagadarikar et al. in [17]—namely the coded aperture snapshot spectral imager (CASSI)—which modulates and disperses the spectral density source before it impinges into a camera detector. Furthermore, different variants of the CASSI has been proposed such as the dual disperser CASSI (DD-CASSI) [18], 3D-CASSI [16], and colored-CASSI [19]. In addition, a multi-sensor CSI architecture was recently used to recover high-resolution spectral images from hyperspectral and multispectral compressive samples [20].

A naïve approach to label SIs from compressive measurements would involve the reconstruction of the spectral image, and subsequently, the application of a pixel-based classifier to the recovered image.

In 2014, Ramirez et al. proposed a classification method that uses optimal coded aperture patterns in CASSI. The entries of the coded apertures are optimized according to the restricted isometry property (RIP). The classification is performed using a support vector machine (SVM), and nearest neighbor classifier [21]. Since this method does not consider the spatial contextual information embedded in spectral images, its performance is severely degraded at higher compression rates. Moreover, the CS classification algorithm exhibits high computational costs.

Recently, two spectral image classification approaches from multi-sensor compressive measurements have been proposed in [22], [23]. However, these approaches do not include an acquisition procedure that considers the spatial contextual information in spectral images to prevent the salt and pepper noise in classification maps. On the other hand, adaptive acquisition schemes have been proposed to prevent saturation in the detector [24], [25] and to improve the quality of image reconstruction [26]. A classification approach in the compressive domain is reported in [27] that develops an adaptive classification scheme from DD-CASSI measurements [18]. This approach consists of an adaptive coded aperture design method based on a Bayesian classifier and a design framework to obtain classification features. This technique has been developed to solve a particular problem, and its extension to other CSI acquisition architectures has become a challenging issue. As far as we know, no other CSI systems, i.e., 3D-CASSI and C-CASSI, are performing adaptive CSI acquisition for SI classification.

This paper focuses on developing an adaptive acquisition method to classify SI directly from multi-sensor compressive measurements. More precisely, the proposed sensing approach considers the spatial contextual information in spectral images to adaptively capture the relevant spatial-spectral information of the scene of interest, thus improving the classification performance. Moreover, the proposed approach is developed for dual-arm systems equipped with two CSI optical architectures: 3D-CASSI and C-CASSI. The contributions of this paper are three-fold.

- 1) We develop an adaptive acquisition method to capture multi-sensor compressive measurements to classify high-resolution spectral images without resorting to a spectral image reconstruction algorithm. More precisely, the proposed acquisition approach first takes advantage of the spatial contextual information embedded in MS projections to adaptively design the colored coded aperture patterns of a multi-sensor compressive optical system.
- 2) Secondly, we propose an algorithm that extracts classification features from adaptive compressive measurements. This algorithm attains outstanding classification results for a reduced set of multi-sensor compressive projections, and it also remarkably improves the labeling accuracy as the number of compressive snapshots increase.

- 3) The proposed approach outperforms other classification methods from compressive measurements.

This paper is organized as follows. Section II presents the observation model, which describes the optical model comprised of a dual-arm C-CASSI and 3D-CASSI system. Section III describes the proposed adaptive CSI classification scheme. Section IV outlines the results using the adaptive approach using synthetic and experimental hyperspectral datasets. Conclusions are summarized in section V.

II. OBSERVATION MODEL

Notice that the proposed classification method is applied to compressive measurements captured by dual-arm systems based on the 3D-CASSI [16] and the C-CASSI [19] optical architectures. Therefore, we start by presenting a brief description of the CASSI-based sensors. Afterward, we introduce the proposed dual-arm architecture comprised of a high spatial resolution system and a high spectral resolution system. A specific description of the dual-arm-based 3D-CASSI and C-CASSI is presented in this section.

A. CASSI based optical architectures

In order to describe the CASSI based optical systems, consider $f_0(x, y, \lambda)$ as the spectral density source associated with the scene of interest, where (x, y) represents the spatial location and λ denotes a particular wavelength. In particular, CASSI based architectures encode the spectral density source by including a coded aperture that typically consists of an array of optical filters. In this sense, the encoded spectral density can be described as

$$f_1(x, y, \lambda) = f_0(x, y, \lambda)t(x, y, \lambda), \quad (1)$$

where $t(x, y, \lambda)$ is a function that describes the spatial-spectral encoding operation performed by the coded aperture. Then, the encoded field is spectrally shifted by a dispersive element. Thus, the encoded-and-shifted spectral field is given by

$$f_2(x, y, \lambda) = \iiint f_1(x', y', \lambda') \times h(x - \psi(\lambda) - x', y - y', \lambda - \lambda') dx' dy' d\lambda', \quad (2)$$

where $h(x, y, \lambda)$ is the impulse response of the optical system, and $\psi(\lambda)$ describes the shifting effect induced by the optical dispersive element. In general, the sensor's impulse response is assumed as the Dirac delta function, i.e. $h(x, y, \lambda) = \delta(x, y, \lambda)$. Furthermore, it is worth noting that the 3-D CASSI system does not contain the dispersive element [16], therefore, for this system, it is assumed that $\psi(\lambda) = 0$. Subsequently, the encoded-and-dispersed spectral field is integrated over the spectral sensitivity of the detector plane. In this regard, the projected plane can be expressed as

$$g(x, y) = \int_{\Lambda} f_2(x, y, \lambda). \quad (3)$$

Moreover, let Δ be the pixel width of the camera detector; thus, the intensity captured by the detector at the discrete coordinate (i, j) can be described as

$$(\mathbf{Y})_{(i,j)} = \int_{\Delta} \int_{\Delta} g(x, y) \text{rect} \left(\frac{x}{\Delta} - i, \frac{y}{\Delta} - j \right) dx dy, \quad (4)$$

where $\text{rect} \left(\frac{x}{\Delta} - i, \frac{y}{\Delta} - j \right)$ represents the area covered by the detector pixel at the spatial location (i, j) . The expression of the intensity captured by a CASSI based detector at the location (i, j) is described by the discrete model shown as follows

$$(\mathbf{Y})_{i,j} = \sum_{\ell=0}^{L-1} \mathcal{F}_{i,(j-c\ell),\ell} \mathcal{T}_{i,(j-c\ell),\ell} + \eta_{i,j}, \quad (5)$$

where $\mathcal{F} \in \mathbb{R}^{M \times N \times L}$ is a discrete version of the input spectral image with dimensions of $M \times N$ pixels and L spectral bands. Every element of the input spectral image is denoted as $\mathcal{F}_{i,j,\ell}$ with ℓ as the spectral index. Notice that $\mathcal{T} \in \{0, 1\}^{M \times N \times L}$ is a binary cube with entries $\mathcal{T}_{i,j,\ell}$ that describes the spatial-spectral encoding operation performed by the coded aperture; $\eta_{i,j}$ is the noise inherent to the sampling system; and c is a factor related to the spectral shifting operation induced by the dispersive element which typically is set to $c = 1$ for C-CASSI.

Multiple snapshots are frequently required to reconstruct a reliable version of the discrete spectral image from CASSI compressive measurements, where each snapshot is captured using a different encoding pattern. Hence, the intensity captured at the coordinate (i, j) and the k^{th} snapshot can be given by

$$(\mathbf{Y})_{i,j}^k = \sum_{\ell=0}^{L-1} \mathcal{F}_{i,(j-c\ell),\ell} \mathcal{T}_{i,(j-c\ell),\ell}^k + \eta_{i,j}^k, \quad (6)$$

for $k = 0, \dots, K$, where K is the number of captured snapshots and $\mathcal{T}^k \in \{0, 1\}^{M \times N \times L}$ is the binary model of a distinct encoding pattern used at the k^{th} snapshot. Finally, the vectorized CASSI compressive measurements can be obtained as

$$\mathbf{y} = \mathbf{H}\mathbf{f} + \boldsymbol{\eta}, \quad (7)$$

where $\mathbf{y} \in \mathbb{R}^{M(N+cL-c)K}$ is the vector that contains the set of compressive samples; $\mathbf{H} \in \mathbb{R}^{M(N+cL-c)K \times (MNL)}$ denotes the matrix that describes the acquisition process including the encoding operation performed by the coded aperture, the dispersive element effect, and the projection onto the camera detector; $\mathbf{f} \in \mathbb{R}^{MNL}$ is the input spectral image in vector form; and $\boldsymbol{\eta} \in \mathbb{R}^{M(N+cL-c)K}$ is the additive noise vector.

B. Dual-arm architecture based on C-CASSI

Figure 1 illustrates a representation scheme of the dual-arm architecture based on the C-CASSI optical system. As can be seen, the beamsplitter divides the input spectral density into two optical paths. The upper arm, which is referred to as the multispectral arm, is a C-CASSI sensor that contains two relay lenses, a dispersive element, a high-spatial-resolution coded aperture, and a high-spatial-resolution detector plane. On the other hand, the bottom arm called the hyperspectral arm is a

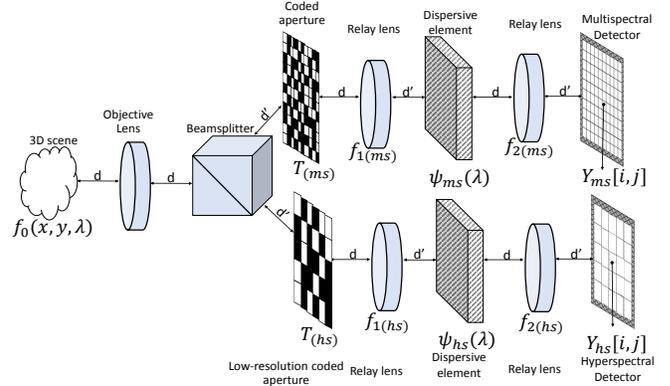


Fig. 1: Schematic of the dual-arm architecture based on the colored-CASSI optical system with a multispectral sensor and a hyperspectral sensor.

C-CASSI system equipped with a low-spatial-resolution coded aperture and a low-spatial-resolution camera detector.

This paper proposes a spectral image classification method from compressive measurements captured by the dual-arm system. More precisely, the measurements captured by the dual-arm system are adaptively obtained by considering the spatial contextual information of the scene of interest with the aim of improving the classification performance using a reduced set of measurements. To this end, we also propose two algorithms that adaptively design the encoding patterns of both the multispectral arm coded aperture and the hyperspectral arm code aperture. Therefore, the multispectral arm and hyperspectral arm acquisition processes for the C-CASSI is described next.

C-CASSI multispectral arm: This branch captures high-spatial but low-spectral resolution compressive samples. More precisely, the compressive measurements obtained by the C-CASSI multispectral arm can be modeled as

$$\bar{\mathbf{y}}_{(\text{ms})} = \bar{\mathbf{H}}_{(\text{ms})} \mathbf{D}_{(\text{ms})} \mathbf{f} + \boldsymbol{\eta}_{(\text{ms})} \quad (8)$$

where $\bar{\mathbf{H}}_{(\text{ms})}$ is the measurement matrix of the compressive multispectral arm and $\mathbf{D}_{(\text{ms})} \in \mathbb{R}^{\frac{MNL}{q} \times MNL}$ is the downsampling matrix of that reduces the number of spectral bands with q as the spectral decimation factor. The measurement matrix that describes the acquisition system at the k^{th} snapshot $\bar{\mathbf{H}}_{(\text{ms})}^k$ can be defined as

$$\bar{\mathbf{H}}_{(\text{ms})}^k = \begin{pmatrix} \mathbf{T}_{(\text{ms})}^{(k,1)} & \mathbf{0}_{M(1) \times MN} & \cdots & \mathbf{0}_{M(L-1) \times MN} \\ \vdots & \mathbf{T}_{(\text{ms})}^{(k,2)} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{M(L-1) \times MN} & \mathbf{0}_{M(L-2) \times MN} & \cdots & \mathbf{T}_{(\text{ms})}^{(k, \frac{L}{q})} \end{pmatrix}. \quad (9)$$

where $\mathbf{T}_{(\text{ms})}^{(k,\ell)} = \text{diag}(\mathcal{T}_{::,\ell}^{(\text{ms})})$ is the diagonalized version of the ℓ^{th} band extracted from the binary model that describes the coded aperture used by the compressive multispectral arm at the k^{th} snapshot, and $\mathbf{0}_{M \times N}$ is a matrix with dimensions $M \times N$ whose entries are zero. It is worth noting that the proposed approach adaptively designs the measurement matrices, and consequently the coded aperture patterns, to improve the image classification performance.

On the other hand, the multispectral decimation matrix can be defined by $\mathbf{D}_{(ms)} = \mathbf{D}_\lambda \otimes \mathbf{I}_{MN}$, being $\mathbf{D}_\lambda \in \mathbb{R}^{\frac{L}{q} \times L}$ the matrix that averages the spectral bands of the original datacube, \otimes denotes the Kronecker product, and $\mathbf{I}_{MN} \in \mathbb{R}^{MN \times MN}$ is an identity matrix. Notice that the multispectral image can be obtained as $\mathbf{f}_{(ms)} = \mathbf{D}_{(ms)}\mathbf{f}$, and $\boldsymbol{\eta}_{(ms)}$ is the additive noise vector that contaminates the compressive multispectral samples.

C-CASSI hyperspectral arm: This arm acquires low-spatial resolution compressive measurements with rich spectral information. In particular, the compressive samples captured by the hyperspectral arm can be succinctly described as

$$\bar{\mathbf{y}}_{(hs)} = \bar{\mathbf{H}}_{(hs)}\mathbf{D}_{(hs)}\mathbf{f} + \boldsymbol{\eta}_{(hs)}, \quad (10)$$

where $\bar{\mathbf{H}}_{(hs)}$ is the measurement matrix of the compressive hyperspectral arm, $\mathbf{D}_{(hs)} \in \mathbb{R}^{\frac{MNL}{p^2} \times MNL}$ is the downsampling matrix with p as the spatial decimation factor, and $\boldsymbol{\eta}_{(hs)}$ is the additive noise vector. In this case, the measurement matrix that characterizes the compressive acquisition process at the k^{th} snapshot can be modeled as

$$\bar{\mathbf{H}}_{(hs)}^k = \begin{pmatrix} \mathbf{T}_{(hs)}^{(k,1)} & \mathbf{0}_{\frac{M}{p}(1) \times \frac{MN}{p^2}} & \cdots & \mathbf{0}_{\frac{M}{p}(L-1) \times \frac{MN}{p^2}} \\ \mathbf{0} & \mathbf{T}_{(hs)}^{(k,2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{\frac{M}{p}(L-1) \times \frac{MN}{p^2}} & \mathbf{0}_{\frac{M}{p}(L-2) \times \frac{MN}{p^2}} & \cdots & \mathbf{T}_{(hs)}^{(k,L)} \end{pmatrix} \quad (11)$$

where $\mathbf{T}_{(hs)}^{(k,\ell)} = \text{diag}(\mathcal{T}_{::,\ell}^{(hs)})$ is the diagonalized version of the ℓ^{th} band extracted from the binary model that characterizes the coded aperture used by the compressive hyperspectral arm at the k^{th} snapshot. A detailed description of $\mathbf{D}_{(hs)}$ is presented below.

- The spatial decimation matrix $\mathbf{D}_{(hs)} \in \mathbb{R}^{\frac{MNL}{p^2} \times MNL}$ can be described as $\mathbf{D}_{(hs)} = \mathbf{I}_{(hs)} \otimes \mathbf{S}_{(hs)}$ where $\mathbf{I}_{(hs)} \in \mathbb{R}^{L \times L}$ is an identity matrix, and $\mathbf{S}_{(hs)} \in \mathbb{R}^{\frac{MN}{p^2} \times MN}$, it is a matrix that degrades spatially each band, it is given by $\mathbf{S}_{(hs)} = \mathbf{D}_c \mathbf{D}_r$.
- The matrix $\mathbf{D}_c \in \mathbb{R}^{\frac{MN}{p^2} \times \frac{MN}{p^2}}$ decimates the columns of each spectral band. In more detail, $\mathbf{D}_c = \mathbf{I}_c \otimes \mathbf{G}^T$ where $\mathbf{G}^T \in \mathbb{R}^{\frac{N}{p} \times N}$ is given by $G_l = \boldsymbol{\Theta}_c^{(l)} \mathbf{d}_c$ where $l = \{0, 1, \dots, \frac{N}{p} - 1\}$ corresponds to the index of the columns of matrix \mathbf{G} , $\mathbf{d}_c \in \mathbb{R}^N$ is a vector that selects of columns, the position of the ones are determined by $\frac{iN}{p}$ with $i = \{0, 1, \dots, p-1\}$ and $\boldsymbol{\Theta}_c^{(l)} \in \mathbb{R}^{N \times N}$ is a permutation matrix is given by,

$$\boldsymbol{\Theta}_c = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \quad (12)$$

- The matrix $\mathbf{D}_r \in \mathbb{R}^{\frac{MN}{p} \times MN}$ decimate the rows of each spectral band, which is given by $\mathbf{D}_r = \mathbf{I}_r \otimes \mathbf{d}_r$ where $\mathbf{d}_r = \mathbf{1}^T$ is a vector of all ones such as $\mathbf{d}_r \in \mathbb{R}^p$.

C. Dual-arm architecture based on 3D-CASSI

The compressive measurements of the 3D-CASSI preserve better the spatial distribution of the acquired projection than the C-CASSI compressive measurements. The compressive measurements of 3D-CASSI are given by Eq. (6) when $c = 0$, denoting the absence of dispersion in the system. In the following are presented the measurement matrix of the multispectral arm and hyperspectral arm.

3D-CASSI multispectral arm: The compressive samples captured by the 3D-CASSI multispectral arm can be also expressed by (8), where the measurement matrix representing the acquisition system at the k^{th} snapshot is defined as

$$\bar{\mathbf{H}}_{(ms)}^k = \left[\mathbf{T}_{(ms)}^{(k,1)}, \mathbf{T}_{(ms)}^{(k,2)}, \dots, \mathbf{T}_{(ms)}^{(k,\frac{L}{q})} \right]. \quad (13)$$

3D-CASSI hyperspectral arm: The compressive measurements obtained by the 3D-CASSI hyperspectral branch can be represented by (10), where the sampling matrix at the k^{th} snapshot can be described as

$$\bar{\mathbf{H}}_{(hs)}^k = \left[\mathbf{T}_{(hs)}^{(k,1)}, \mathbf{T}_{(hs)}^{(k,2)}, \dots, \mathbf{T}_{(hs)}^{(k,\frac{L}{q})} \right]. \quad (14)$$

III. DESIGN OF ADAPTIVE ACQUISITION

To perform the adaptive acquisition is required to compute the contextual information of the scene. The contextual spatial information is used to design the coded apertures. This section describes the hardware and processing algorithm that computes the contextual spatial information and the adaptive coded aperture design. In particular, Fig. 2 shows the optical and the processing algorithm that allows attaining the compressive measurements and the classification maps, respectively.

A. Computation of the contextual information

Grayscale image estimation: The grayscale image is estimated to design the coded aperture by exploiting the spatial resolution in the multispectral arm. The intuition is to reduce the dispersion in the captured measurement. The acquisition of the grayscale is different depending on the CSI architecture. For the 3D-CASSI, the colored coded aperture of the first snapshot in the multispectral arm is set to an all-ones matrix $\mathcal{T}_{i,j,\ell,0}^{(ms)} = 1$, exploiting the lack of disperser. This structure has the same effect as removing the coded aperture. For the C-CASSI, the dispersion is reduced using a colored coded aperture with bandpass filter $\mathcal{T}_{i,j,m,0}^{(ms)} = 1$ in all the spatial position, where m denotes the middle wavelengths. The bandpass filter samples the central wavelengths. The purpose of this step is to attain the first multispectral compressive measurement $\mathbf{Y}_{(ms)}^0$ with lower dispersion or approximately a grayscale image. Equation (6) denotes the acquisition of first and successive compressive measurements. In C-CASSI, the compressive measurements are cropped to retain the scene's spatial size. The contextual spatial information is obtained from the estimated grayscale images using two processes, high-pass filtering, and quantization.

High-pass filter: The high-pass filtering process obtains the components of the high frequency of the estimated grayscale

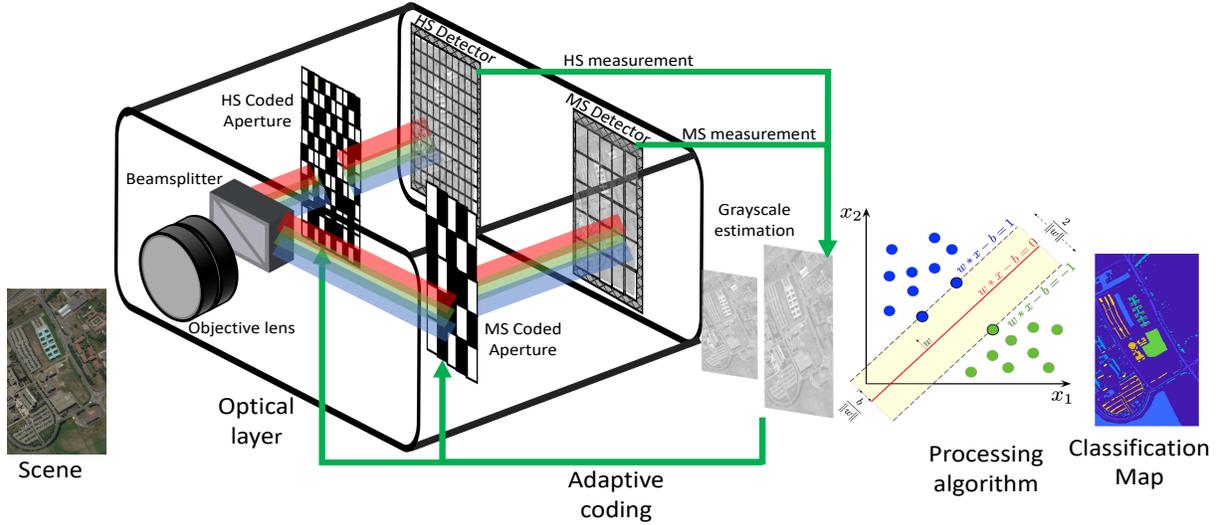


Fig. 2: Sketch of the optical and processing algorithm of the 3D-CASSI. In the optical layer, the light of the scene enters through the objective lens. After that, the incoming light is divided by the beamsplitter. Thus, the beam in each arm is encoded by the coded aperture and multiplexed onto the detector. The compressive measurements are used to estimate the grayscale image of the scene. The subsequently coded apertures are computed using the contextual spatial information of the grayscale image estimation. At the end of the sensing process, the processing algorithm computes the classification map.

Algorithm 1 Classification using adaptive coded aperture design, and exploiting spatial contextual information in a multi-sensor CSI.

Input: $\mathbf{Y}_{(ms)}^0$ ▷ Inputs of the algorithm.
Output: \mathbf{C}_{test} . ▷ Output of the algorithm.

- 1: **function** CLASSIFICATION USING ADAPTIVE CODING($\mathbf{Y}_{(ms)}^0$)
- 2: $\bar{\mathbf{Y}}_{(ms)}^0 \leftarrow$ **High-pass filter**($\mathbf{Y}_{(ms)}^0$) ▷ High-pass filtering of grayscale image.
- 3: $\mathbf{\Gamma}_{(ms)} \leftarrow$ **Quantization**($\bar{\mathbf{Y}}_{(ms)}^0$) ▷ Quantization of grayscale image.
- 4: $\mathbf{M}_{(ms)} \leftarrow$ $\mathbf{E}_{(ms)} \mathbf{I}_{(ms)} \otimes \mathbf{1}_{(ms)}$ ▷ Matched filter in MS arm.
- 5: $\mathbf{M}_{(hs)} \leftarrow$ $\mathbf{E}_{(hs)} \mathbf{I}_{(hs)} \otimes \mathbf{1}_{(hs)}$ ▷ Matched filter in HS arm.
- 6: $\Xi_{(ms)} \leftarrow$ **Adaptive coding**($\mathbf{M}_{(ms)}, \mathbf{\Gamma}_{(ms)}$) ▷ Computes the position for each class solving Eq. (15).
- 7: $\Xi_{(hs)} \leftarrow$ **Adaptive coding**($\mathbf{M}_{(hs)}, \mathbf{\Gamma}_{(hs)}$) ▷ Computes the position for each class solving Eq. (17).
- 8: $[\mathbf{Y}_{(ms)}^1 \dots \bar{\mathbf{Y}}_{(ms)}^{W-1}] \leftarrow$ MULTISPECTRAL-ARM($\mathbf{\Gamma}_{(ms)}, \Xi_{(ms)}$)
- 9: $[\mathbf{Y}_{I(hs)}^1 \dots \bar{\mathbf{Y}}_{I(hs)}^{K-1}] \leftarrow$ HYPERSPECTRAL-ARM($\mathbf{\Gamma}_{(hs)}, \Xi_{(hs)}$)
- 10: $\mathbf{R} \leftarrow$ $[\text{vec}(\bar{\mathbf{Y}}_{(ms)}^0), \dots, \text{vec}(\bar{\mathbf{Y}}_{(ms)}^{W-1}), \text{vec}(\bar{\mathbf{Y}}_{I(hs)}^0), \dots, \text{vec}(\bar{\mathbf{Y}}_{I(hs)}^{K-1})]^T$ ▷ Spatio-spectral feature extraction.
- 11: $\mathbf{C}_{training} \leftarrow$ **SVM**_{training}(\mathbf{R}) ▷ Classification labels using training data.
- 12: $\mathbf{C}_{test} \leftarrow$ **SVM**_{test}(\mathbf{R}) ▷ Classification labels using test data.
- 13: **return** \mathbf{C}_{test} ▷ Output of the algorithm.

image $\bar{\mathbf{Y}}_{(ms)}^0$, i.e., detecting the edge of the estimated grayscale image. The filtering is computed using $\bar{\mathbf{Y}}_{(ms)}^0(i,j) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} h_{(i-u,j-v)} \mathbf{Y}_{(ms)}^0(i,j)$, where $\bar{\mathbf{Y}}_{(ms)}^0(i,j)$ is the output image of the high pass filtering, $h_{(i,j)}$ is the filter impulse response, $\mathbf{Y}_{(ms)}^0(i,j)$ is the compressive measurement.

Quantization: In the quantization stage, the filtered image $(\bar{\mathbf{Y}}_{(ms)}^0)_{(i,j)}$ is quantized according to the number of classes in the scene n_c . Due to our approach is supervised n_c is known, then it is assumed that the step size in the quantizer $\gamma = n_c$. The typical quantizer can be seen as segmentation method $(\mathbf{\Gamma})_{(ms)(i,j)} = \gamma \left\lfloor \frac{(\bar{\mathbf{Y}}_{(ms)}^0)_{(i,j)}}{\gamma} + \frac{1}{2} \right\rfloor$, where $\lfloor \cdot \rfloor$ is the floor function, γ is the quantizer's step size. The matrix

$\mathbf{\Gamma}_{(ms)} \in \{0, \gamma - 1\}^{M \times N}$ provides the prior spatial contextual information to the algorithm. The quantization matrix in the hyperspectral arm $\mathbf{\Gamma}_{(hs)} \in \{0, \gamma - 1\}^{\frac{M}{p} \times \frac{N}{p}}$ that is $\frac{1}{p}$ times the size compressive measurements $\mathbf{Y}_{(ms)}$.

B. Matched filter

Once it is computed the contextual spatial information; the coded apertures are designed using the matched filter. The Matched filter is computed by $\mathbf{M} = \mathbf{E}\mathbf{B}$, where \mathbf{E} is the average of the training samples of each class, and \mathbf{B} are the designed complementary filters. The complementary filters are a set of filters that adds to an all-pass filter, avoiding the spectral redundancy of compressive measurements. The matched filter computes the spectral response between the designed

complementary filters and the average training samples of each class.

Matched filter in the multispectral arm: In particular, the matched filter in the multispectral arm is defined as $\mathbf{M}_{(ms)} = \mathbf{E}_{(ms)}\mathbf{B}_{(ms)}$, such that $\mathbf{M} \in \mathbb{R}^{\gamma \times n_{(ms)}}$, where it is assumed that $\gamma = n_c$, and the number of filters $n_{(ms)} = W$, with W the number of snapshots in the multispectral arm. Notice that $\mathbf{E}_{(ms)} \in \mathbb{R}^{\gamma \times \frac{L}{q}}$, where $\frac{L}{q}$ is the number of bands in the multispectral arm, and each column in $\mathbf{E}_{(ms)} = [\mathbf{e}_{(ms):,0}, \mathbf{e}_{(ms):,1}, \dots, \mathbf{e}_{(ms):,\gamma-1}]^T$ is the average of the training samples for each of the classes. In addition, the matrix $\mathbf{B}_{(ms)} \in \{0, 1\}^{\frac{L}{q} \times n_{(ms)}}$ are the designed filters such that $\mathbf{B}_{(ms)} = [\mathbf{b}_{(ms):,0}, \mathbf{b}_{(ms):,1}, \dots, \mathbf{b}_{(ms):,n_f-1}]$, where each column in $\mathbf{B}_{(ms)}$ is a part of the complementary filter with filter's width $\alpha_{(ms)} = \frac{L}{qW}$. The effect of a complementary filter is sampled a different part of the spectral signature at each snapshot. The designed filter is $\mathbf{B}_{(ms)} = \mathbf{I}_{(ms)} \otimes \mathbf{1}_{(ms)}$ where $\mathbf{1}_{(ms)}$ is all-ones vector such that $\mathbf{1}_{(ms)} \in \mathbb{R}^{\alpha_{(ms)} \times 1}$, $\mathbf{I}_{(ms)} \in \mathbb{R}^{W \times W}$ is the identity matrix, and \otimes is the Kronecker product.

Matched filter in the hyperspectral arm: The matched filter in the hyperspectral arm is denoted as $\mathbf{M}_{(hs)} = \mathbf{E}_{(hs)}\mathbf{B}_{(hs)}$, such that $\mathbf{M}_{(hs)} \in \mathbb{R}^{\gamma \times n_{(hs)}}$, where it is assumed that $\gamma = n_c$, and the number of filters $n_{(hs)} = K$, with K the number of snapshots in the hyperspectral arm. Notice that $\mathbf{E}_{(hs)} \in \mathbb{R}^{\gamma \times L}$, where L is the number of bands in the hyperspectral arm, and each column in $\mathbf{E}_{(hs)} = [\mathbf{e}_{(hs):,0}, \mathbf{e}_{(hs):,1}, \dots, \mathbf{e}_{(hs):,\gamma-1}]^T$ is the average of the training samples for each of the classes. Moreover, the matrix $\mathbf{B}_{(hs)} \in \{0, 1\}^{L \times n_{(hs)}}$ is the designed filter such that $\mathbf{B}_{(hs)} = [\mathbf{b}_{(hs):,0}, \mathbf{b}_{(hs):,1}, \dots, \mathbf{b}_{(hs):,n_f-1}]$, where each column in $\mathbf{B}_{(hs)}$ is a complementary filter with filter's width $\alpha_{(hs)} = \frac{L}{K}$, which means that at each snapshot is capture a different wavelength. The designed filter is $\mathbf{B}_{(hs)} = \mathbf{I}_{(hs)} \otimes \mathbf{1}_{(hs)}$ where $\mathbf{1}_{(hs)}$ is all-ones vector such that $\mathbf{1}_{(hs)} \in \mathbb{R}^{\alpha_{(hs)} \times 1}$, $\mathbf{I}_{(hs)} \in \mathbb{R}^{K \times K}$ is the identity matrix, and \otimes is the Kronecker product. In the successive stage the adaptive coding is performed from the previous matched filter.

C. Adaptive coded aperture design

To adaptive design the coded aperture entries, a sorting process is performed over the filter to determine the order of sampling according to the best spectral response, see Eq. (15) and Eq. (17).

Adaptive coding in the multispectral arm: To design the coded aperture is necessary to know the position of the filters at each snapshot, which can be computed as follows

$$\Xi_{(ms)(:, \Gamma_{(ms)i,j})} = \text{sort}(\mathbf{M}_{(ms)}^T(:, \Gamma_{(ms)i,j})) \quad (15)$$

where $\text{sort}(\cdot)$ returns the sorted positions in descending order of the filters at each class. Notice that $\Xi_{(ms)} \in \{0, \dots, W-1\}^{W \times \gamma}$ and the index of the filter is denoted by $\Xi_{(ms)k, \Gamma_{(ms)i,j}}$, where k indexes the snapshots and $\Gamma_{(ms)i,j}$ is the quantization matrix that index of the classes for each entry. The design of the coded aperture in the multispectral arm is giving by

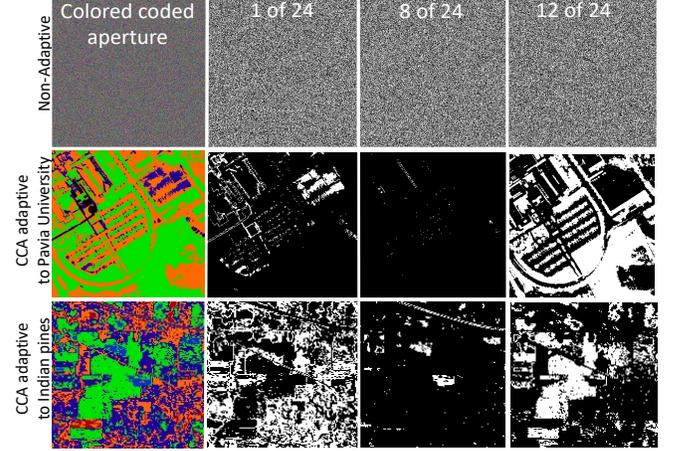


Fig. 3: Comparison between non-adaptive and adaptive colored coded apertures when $W = 6$. The top row shows the non-adaptive random colored coded aperture. The second and third rows depict the adaptive colored coded apertures for the Pavia University and Indian Pines, respectively. Notice the high correlation of the designed codes with the corresponding scene. (First column) Denote the false color representation of the colored coded apertures; (second column) show three slides of the 3D block-unblock representation.

$$\mathcal{T}_{i,j,;k}^{(ms)} = \mathbf{b}_{(ms):, \Xi_{(ms)k, \Gamma_{(ms)i,j}}} \quad (16)$$

where $\mathbf{b}_{(ms):,v}$ is the each designed filter in the v^{th} class. Notice that the designed coded aperture in the ℓ^{th} spectral band for the w^{th} multispectral snapshot can be reorganize as follows $\mathbf{T}_{(ms)}^{(w, \ell)} = \text{diag}(\mathcal{T}_{:, \ell, w}^{(ms)})$. In particular, the $\frac{L}{q}$ coded apertures in $\mathbf{T}_{(ms)}^w$ for the w^{th} snapshot are rearranged to yield the sensing matrices $\bar{\mathbf{H}}_{(ms)}^w$ according to Eq. (9). The random and designed colored coded apertures are depicted in Fig. 3.

Adaptive coding in the hyperspectral arm: In the same manner, the design of the coded apertures for the hyperspectral arm requires the location of the filters at each snapshot, which is denoted by

$$\Xi_{(hs)(:, j)} = \text{sort}(\mathbf{M}_{(hs)}^T(:, \Gamma_{(hs)i,j})) \quad (17)$$

where $\text{sort}(\cdot)$ returns the sorted positions in descending order of the filters at each class, the matrix is $\Xi_{(hs)} \in \{0, \dots, K-1\}^{K \times \gamma}$, and the quantization matrix in the hyperspectral arm $\Gamma_{(hs)}$. The index of each filter is denoted by $\Xi_{(hs)k, \Gamma_{(hs)i,j}}$, where k is the index of hyperspectral snapshots and $\Gamma_{(hs)i,j}$ is the quantization matrix that index of the classes. The design of the coded aperture in the hyperspectral arm is giving by

$$\mathcal{T}_{i,j,;k}^{(hs)} = \mathbf{b}_{(hs):, \Xi_{(hs)k, \Gamma_{(hs)i,j}}} \quad (18)$$

where $\mathbf{b}_{(hs):,v}$ is each designed filter in the v^{th} class. Notice that the coded aperture in the ℓ^{th} band for the k^{th} hyperspectral snapshot is rearranged as $\mathbf{T}_{(hs)}^{(k, \ell)} = \text{diag}(\mathcal{T}_{:, \ell, k}^{(hs)})$. In detail, the L coded apertures in $\mathbf{T}_{(hs)}^k$ for the k^{th} snapshot can be rearranged as the sensing matrix $\bar{\mathbf{H}}_{(hs)}^k$ using Eq.(11). Once, the coded are designed the compressive measurements are captured.

D. Acquisition of compressive measurements

In the fourth stage, each arm captures a different type of compressive measurement. At each snapshot, the multispectral compressive measurements $\mathbf{y}^k_{(hs)} = \bar{\mathbf{H}}^k_{(hs)} \mathbf{D}_{(hs)} \mathbf{f} + \boldsymbol{\eta}^k_{(hs)}$ are captured, and the hyperspectral compressive measurements $\mathbf{y}^w_{(ms)} = \bar{\mathbf{H}}^w_{(ms)} \mathbf{D}_{(ms)} \mathbf{f} + \boldsymbol{\eta}^w_{(ms)}$ are captured. Figure 4 depicts an example of the compressive measurements of the multispectral and hyperspectral arms. Notice that when the coded apertures are spatially and spectrally complementary but randomly distributed, a salt and pepper noise is produced over the compressive measurements (see Fig. 4(a), and 4(b)). In contrast, when the coded apertures are designed the salt pepper noise is prevented (see Fig. 4(c), and 4(d)). Nevertheless, in order to include the spatial contextual information in the extracted features, a denoising stage is required. The following subsection explains the denoising method, and the extraction of features.

E. Denoising and Feature extraction

In the denoising stage, a median filter is applied over the compressive multispectral, and hyperspectral measurements $Y_{i,j,k}^{(hs)}$, and $Y_{i,j,w}^{(ms)}$ to enhance the spatial contextual information. To compute the output of the median filter, an odd number of samples are sorted, and the middle value is used as the output of the filter. If the filter length is $\iota = 2\beta + 1$, the filtering procedure is denoted as $\bar{Y}_{i,j,k}^{(hs)} = \mathbf{MED}[Y_{i,j-\beta,k}^{(hs)}, \dots, Y_{i,j,k}^{(hs)} \dots Y_{i,j+\beta,k}^{(hs)}]$, and $\bar{Y}_{i,j,k}^{(ms)} = \mathbf{MED}[Y_{i,j-\beta,w}^{(ms)}, \dots, Y_{i,j,w}^{(ms)} \dots Y_{i,j+\beta,w}^{(ms)}]$, where $\bar{Y}_{i,j,w}^{(ms)}$, and $\bar{Y}_{i,j,k}^{(hs)}$ are the (i, j) th multispectral, and hyperspectral output, respectively. The median filter with window's size $[7 \times 7]$ is performed to promote the uniformity in both compressive measurements $Y_{i,j,k}^{(hs)}$, $Y_{i,j,w}^{(ms)}$.

In addition, the filtered hyperspectral measurements $\bar{\mathbf{Y}}_{(hs)}^k$ are interpolated using a bilinear interpolator \mathbf{P} to attain the size of the multispectral measurements, resulting in $\bar{\mathbf{Y}}_{I(hs)}^k = \mathbf{P}(\bar{\mathbf{Y}}_{(hs)}^k)$. Using the filtered compressive multispectral measurements $\bar{\mathbf{Y}}_{(ms)}^w$, and the filtered-interpolated compressive hyperspectral measurements $\bar{\mathbf{Y}}_{I(hs)}^k$, the features are extracted in the matrix $\mathbf{R} = [\text{vec}(\bar{\mathbf{Y}}_{(ms)}^0), \dots, \text{vec}(\bar{\mathbf{Y}}_{(ms)}^{W-1}), \text{vec}(\bar{\mathbf{Y}}_{I(hs)}^0), \dots, \text{vec}(\bar{\mathbf{Y}}_{I(hs)}^{K-1})]^T$. Figures 4(e) and 4(f) depict the spatial and spectral features after applying the median filter to the non-adaptive compressive measurements. In contrast, figure 4(g), and 4(h) show the spatial $\bar{\mathbf{Y}}_{(ms)}^w$, and spectral $\bar{\mathbf{Y}}_{I(hs)}^k$ features after applying the median filter to the adaptive compressive measurements. Notice that compressive measurements using the non-adaptive approach attain lower spatial uniformity Fig. 4(e) and Fig. 4(f) than the compressive measurements obtained using the proposed adaptive design.

F. Classification

The classification is performed using as input the extracted features of the previous step. The support vector machine (SVM) classifier [4], using a polynomial kernel computes the

Algorithm 2 Adaptive coding and processing in the hyperspectral-arm.

```

function HYPERSPECTRAL-ARM( $\Gamma_{(hs)}, \Xi_{(ms)}$ )
2:   for  $k \leftarrow 0, K-1$  do
       for  $i \leftarrow 0, \frac{M}{p}$  do
4:     for  $j \leftarrow 0, \frac{N}{p}$  do
            $v \leftarrow \Xi_{(hs)k, \Gamma_{(hs)}(i,j)}$ 
6:        $\mathcal{T}_{i,j,:,k}^{(hs)} \leftarrow \mathbf{b}_{(hs)::v}$ 
            $\mathbf{Y}_{(hs)}^k \leftarrow \mathbf{Acq}(\mathcal{T}_{i,j,:,k}^{(hs)})$  ▷ Hyperspectral
           measurements acquisition using Eq. (6).
8:        $\bar{\mathbf{Y}}_{(hs)}^k \leftarrow \mathbf{Med}(\mathbf{Y}_{(hs)}^k)$  ▷ Compute median
           filtering.
            $\bar{\mathbf{Y}}_{I(hs)}^k \leftarrow \mathbf{P}(\bar{\mathbf{Y}}_{(hs)}^k)$  ▷ Interpolation of
           hyperspectral measurements.
10:    return  $[\bar{\mathbf{Y}}_{I(hs)}^1 \dots \bar{\mathbf{Y}}_{I(hs)}^{K-1}]$ 

```

class labels for training $\mathbf{C}_{training}$, and the class labels for test \mathbf{C}_{test} to each spatial position of the scene.

G. Algorithm

The algorithm 1 summarizes the proposed method. Which has as input the compressive measurement of the multispectral arm $\mathbf{Y}_{(ms)}^0$. Using the $\mathbf{Y}_{(ms)}^0$, the high-pass filter is calculated in step 2. Following, the quantization is computed in step 3. Subsequently, the compressive measurements are captured $\mathbf{Y}_{(ms)}^w$, $\mathbf{Y}_{(hs)}^k$, according to the sampling model in line 8, and line 9. The process to obtain $\mathbf{Y}_{(ms)}^w$, and $\mathbf{Y}_{(hs)}^k$ is shown in algorithm 2, and 3. In the adapting coding process, the sampling matrix is designed using the contextual information of the quantization matrix Γ . The coded apertures $\mathcal{T}_{i,j,:,k}^{(hs)}$, and $\mathcal{T}_{i,j,:,k}^{(ms)}$ are computed in lines 6 of algorithms 2, and 3, respectively. The acquisition of the hyperspectral, and multispectral compressive measurements using the adapted codes is attained in lines 7, and 7 of algorithms 2 and 3. The processing before to attain the features is composed of denoising, and the interpolation. The denoising using median filter is conducted in lines 8 of algorithms 2, and 3, respectively. And, the interpolation is only performed to hyperspectral measurements in line 9 of algorithm 2. Retaking the algorithm 1, the extracted features denoted as matrix \mathbf{R} are computed in line 10. After, the classification is computed in lines 11, and 12. The output of the algorithm is the resulting classification labels \mathbf{C}_{test} .

IV. SIMULATIONS AND RESULTS

This section analyzes the performance of the proposed adaptive technique in terms of spectral image classification accuracy. More precisely, the proposed method is evaluated on synthetic compressive measurements obtained from two real datasets: Pavia University and Indian Pines. Subsequently, the proposed approach is tested on real compressive measurements captured by a laboratory setup.

A. Pavia University dataset

Algorithm 3 Adaptive coding and processing in the multispectral-arm.

```

function MULTISPECTRAL-ARM( $\Gamma_{(ms)}$ ,  $\Xi_{(ms)}$ )
2:   for  $w \leftarrow 1, W - 1$  do
       for  $i \leftarrow 0, M$  do
4:         for  $j \leftarrow 0, N$  do
                $v \leftarrow \Xi_{(ms)w, \Gamma_{(ms)}(i,j)}$ 
6:            $\mathcal{T}_{i,j,:,w}^{(ms)} \leftarrow \mathbf{b}_{(ms):,v}$ 
                $\mathbf{Y}_{(ms)}^w \leftarrow \mathbf{Acq}(\mathcal{T}_{i,j,:,w}^{(ms)})$             $\triangleright$  Acquisition of
               multispectral measurements using Eq. (6).
8:            $\bar{\mathbf{Y}}_{(ms)}^w \leftarrow \mathbf{Med}(\mathbf{Y}_{(ms)}^w)$             $\triangleright$  Compute median
               filtering.
       return  $[\bar{\mathbf{Y}}_{(ms)}^0 \cdots \bar{\mathbf{Y}}_{(ms)}^{W-1}]$ 
    
```

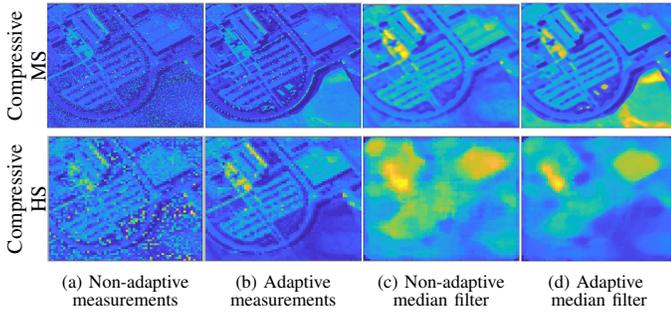


Fig. 4: Denoising of compressive measurements. Multispectral compressive measurements using non-adaptive coded apertures (a), hyperspectral compressive measurements using non-adaptive coded apertures (b), multispectral compressive measurements using adaptive coded apertures (c), hyperspectral compressive measurements using adaptive coded apertures (d). Compressive measurements after median filter with window size $[7 \times 7]$, using non-adaptive coded apertures (e), and (f), using adaptive coded apertures (g) and (h).

The first spectral image was obtained by the reflective optics imaging spectrometer (ROSIS-02) over an urban region of the University of Pavia, Italy [28]. This dataset exhibits dimensions of 640×340 pixels and 103 spectral bands in the wavelength interval from 0.43 to $0.86 \mu\text{m}$. In our simulations, we use a cropped image with a size of $256 \times 256 \times 96$ to efficiently implement the Wavelet transform required by non-adaptive approaches. The ground truth labeling map of $n_c = \gamma = 9$ classes is illustrated in Fig. 5(a), where every label refers to a distinct material in the land cover.

To evaluate the performance of the proposed method, the compressive measurements are first obtained. To this end, the

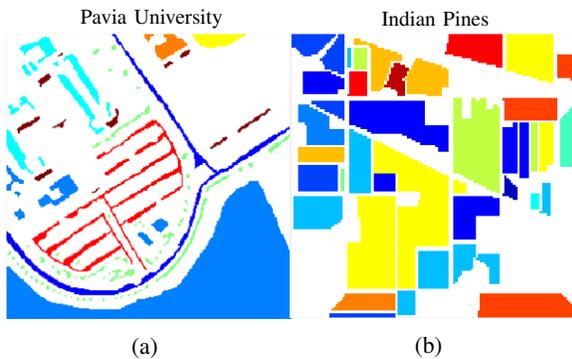


Fig. 5: Ground truth. (a) Pavia University and (b) Indian pines.

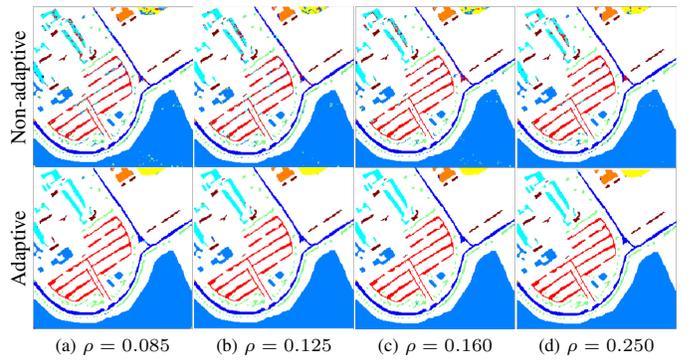


Fig. 6: Pavia University. Labeling maps from 3D-CASSI samples. Non-adaptive approach (a) $\rho = 0.085$, OA:93.32% (b) $\rho = 0.125$, OA:94.89% (c) $\rho = 0.160$, OA:95.05% (d) $\rho = 0.250$, OA:96.26%. Proposed (a) $\rho = 0.085$, OA:99.25% (b) $\rho = 0.125$, OA:99.49% (c) $\rho = 0.160$, OA:99.64% (d) $\rho = 0.250$, OA:99.68%.

HS image is generated by applying a spatial downscaling to the high-resolution spectral image with a decimation factor $p = 4$. In this case, the HS image has dimensions of 64×64 pixels and 96 spectral bands. Then, we simulate the CSI acquisition system to capture the HS compressive measurements, where the compression ratio depends on the number of captured shots. In this work, the proposed approach is evaluated for two CSI acquisition systems: 3D-CASSI [16] and C-CASSI [19]. Furthermore, we use four different compression ratios $\rho = \{0.085, 0.125, 0.160, 0.250\}$ that correspond to $k = \{8, 12, 16, 24\}$ snapshots. On the other hand, the MS image is obtained by applying spectral downsampling to the high-resolution image with a downscaling factor $q = 4$. Hence, the MS image exhibits a size of $256 \times 256 \times 24$. Afterward, the MS compressive measurements are obtained for the various compression ratios that correspond to $w = \{2, 3, 4, 6\}$ shots.

1) *3D-CASSI*: Fig. 6 displays the labeling maps obtained by the proposed adaptive method from 3D-CASSI samples for different compression ratios $\rho = \{0.085, 0.125, 0.160, 0.250\}$. For comparative purposes, the classification maps yielded by the non-adaptive approach [22] are also shown. We use the support vector machine (SVM) method as a supervised classifier with the polynomial kernel of degree $d = 3$. It is clear that the non-adaptive approach attains lower classification values in terms of the overall accuracy (OA) 93.32%, 94.89%, 95.05%, and 96.26% in comparison with the results obtained using the proposed adaptive approach 99.25%, 99.49%, 99.64%, and 99.68%, respectively. In addition, note that the compression ratio of 3D-CASSI and C-CASSI are given by

$$\rho = K \left(\frac{M(N + cL - c)}{p^2 MNL} \right) + W \left(\frac{M(N + \frac{cL}{p} - c)}{MNL} \right) \quad (19)$$

where c is a parameter related to the spectral shifting of the encoded image, k and W denote the number of snapshots captured by the hyperspectral and multispectral arm, respectively. The proposed approach attains OA= 99% using 2 multispectral shots and 8 hyperspectral shots. In contrast, the non-adaptive approach obtains OA= 93% for the same number of snapshots. Notice that there is a significant reduction in the number of shots using the proposed approach.

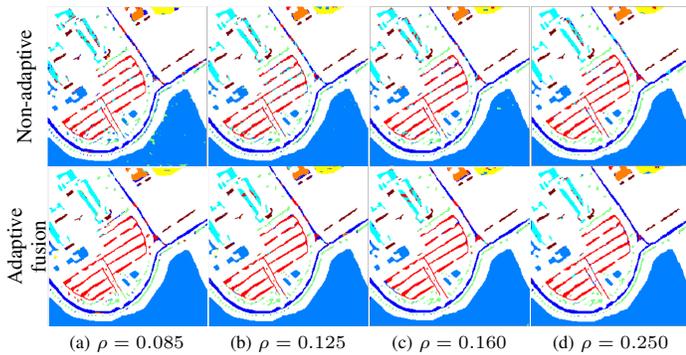


Fig. 7: Pavia University. Labeling maps from C-CASSI samples. Non-adaptive approach (a) $\rho = 0.085$, OA:90.93% (b) $\rho = 0.125$, OA:93.52% (c) $\rho = 0.160$ OA:94.49%, (d) $\rho = 0.250$, AO:96.00%. Proposed (a) $\rho = 0.085$, OA:95.47% (b) $\rho = 0.125$, OA:96.54% (c) $\rho = 0.160$, OA:97.07% (d) $\rho = 0.250$, OA:97.66%.

2) *C-CASSI*: The C-CASSI sensor is a challenging system compared to architectures without dispersion because the contextual information is blurred in compressive projections. Fig. 7 depicts the classification maps obtained from C-CASSI measurements. It also shows a comparison between the non-adaptive C-CASSI and adaptive C-CASSI for the Pavia University dataset. Note that the non-adaptive approach achieves lower classification values in the OA metric 90.93%, 93.52%, 94.49%, and 96.0% against the results obtain utilizing the adaptive approach 95.47%, 96.54%, 97.07%, and 97.66%. There is a reduction in the number of snapshots. The non-adaptive approach attains an accuracy of 90.93% using 2 and 8 multispectral and hyperspectral snapshots, respectively. In contrast, using the same number of snapshots, the adaptive approach obtains an accuracy of 95.47%.

B. Indian Pines dataset

This spectral image was acquired by the airborne visible, infrared imaging spectrometer (AVIRIS) over the Indian Pines, USA [28]. The original size of the Indian Pines dataset is 145×145 pixels and 224 spectral bands in the wavelength range from 0.5 to 2.5 μm . Our experiments use a cropped image with dimensions of $128 \times 128 \times 96$ because non-adaptive classification approaches require dyadic dimensions to efficiently implement the Wavelet transform. Fig. 5(a) shows the ground truth labeling map with $n_c = \gamma = 16$ different classes that identify different crops and building structures.

1) *3D-CASSI*: Fig. 8 shows a comparison of the non-adaptive 3D-CASSI and adaptive 3D-CASSI for the Indian Pines dataset. It is clear that the non-adaptive approach attains low classification values in the OA metric 78.39%, 84.57%, 86.42%, and 90.82% in comparison with the results obtained using the adaptive approach 85.56%, 90.89%, 91.76%, and 92.69% for the compression ratio 0.083, 0.125, 0.16, 0.25. The proposed approach attains 85% of OA in classification using 2 multispectral shots and 8 hyperspectral shots. In contrast, the non-adaptive approach obtains 78.39% with the same number of snapshots. Notice that there is a significant reduction in the number of shots using the proposed approach.

2) *C-CASSI*: Fig. 9 depicts the results using C-CASSI. It depicts a comparison of the non-adaptive C-CASSI and adaptive C-CASSI for the Indian Pines dataset. The test varies the index of multispectral and hyperspectral in the same manner that for 3D-CASSI. Notice that the non-adaptive approach achieves low classification values in the OA metric 87.67%, 88.78%, 90.49%, and 91.02% against the results obtained utilizing the adaptive approach 88.85%, 90.89%, 91.00%, and 92.33%. There is a reduction in the number of snapshots. The non-adaptive approach attains an accuracy of 87.67% using 2 and 8 multispectral and hyperspectral snapshots, respectively. In contrast, using the same number of snapshots, the adaptive approach obtains an accuracy of 88.85%.

C. Experiments with real data

To experimentally prove the benefit of the proposed classification approach the dual-arm C-CASSI was used to acquire compressive measurements in the multispectral and hyperspectral in accordance with [29]. The apparatus is made up of an objective lens (Thorlabs, MAP10100100-A) and a DMD (Texas Instruments, DLI4130VIS-7XGA) to encode the spectral image. The optical setup is composed of two arms. The multispectral arm employs a relay lens (Thorlabs, MAP10100100-A), a dual Amici prism (Shanghai Optics, custom made), and a monochrome sensor (AVT, Stingray F-080B) with 1032×776 pixels and pixel size of $4.65 \times 4.65 \mu\text{m}$. The hyperspectral arm utilizes a 4F-relay system build using two lenses (Thorlabs, AC254-100-A-ML), with a transmission diffraction grating (Thorlabs, GT50-03, 300 grooves/mm, 17.5° groove angle) in the middle, and a monochrome sensor (AVT, Stingray F-080B).

Hereinafter, the evaluation of the performance of the proposed classification approach is presented. One target scene, named Hen is used to evaluate the performance of the computational classification setup. In particular, figure 10 depicts the ground-truth and the RGB version of the scene using the testbed implementation. The test use measurements data of 3D-CASSI, C-CASSI optical setup. A combination of multispectral and hyperspectral snapshots $w = \{2, 3, 4, 6\}$ and $k = \{8, 12, 16, 24\}$, respectively. The spatial resolution of the emulated dataset is 512×512 , and spectral resolution is 96 bands, and $n_c = 4$. To compare, figures 11, and 12 shown the classification map obtained using the non-adaptive and adaptive approach the 3D-CASSI and the adaptive C-CASSI. A significant improvement is shown when the compression ratio is $\rho = 0.085$, 1.43% for the 3D-CASSI, and 18.44% C-CASSI.

V. CONCLUSION

A method that adaptively designs the coded apertures to classify multispectral and hyperspectral compressive measurements of spectral images has been developed. In detail, the approach captures a prior estimation of the grayscale image in the multispectral arm using a coded aperture with all entries with a bandpass filter to reduce the dispersion of the measurement of the first snapshot. The discrete mathematical model to design adaptive colored-coded apertures of the dual-arm

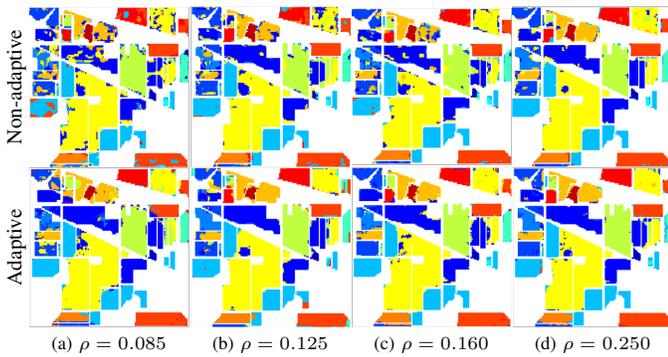


Fig. 8: Indian Pines. Labeling maps from 3D-CASSI measurements. Non-adaptive approach (a) $\rho = 0.085$, OA:78.39%, (b) $\rho = 0.125$, OA:84.57%, (c) $\rho = 0.160$, OA:86.42%, (d) $\rho = 0.250$ OA:90.82%. Proposed (a) $\rho = 0.085$, OA:85.56%, (b) $\rho = 0.125$, OA:90.89%, (c) $\rho = 0.160$, OA:91.76%, (d) $\rho = 0.250$, OA:92.69%.

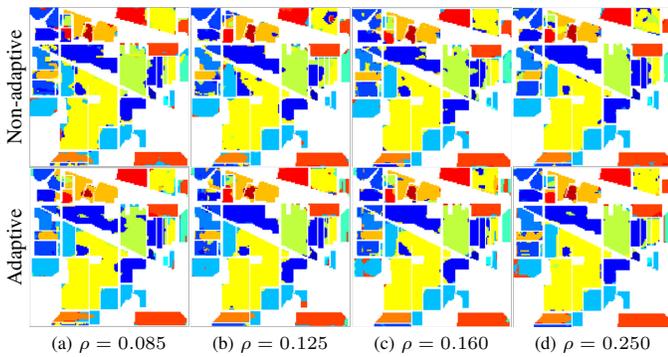


Fig. 9: Indian Pines. Labeling maps from C-CASSI measurements. Non-adaptive approach (a) $\rho = 0.085$, OA:87.67% (b) $\rho = 0.125$, OA:88.78% (c) $\rho = 0.160$, OA:90.49% (d) $\rho = 0.250$, OA:91.02%. Proposed (a) $\rho = 0.085$, OA:88.85% (b) $\rho = 0.125$, OA:90.89% (c) $\rho = 0.160$, OA:91.00% (d) $\rho = 0.25$, AO:92.33%.

CSI is presented. The approach filter, and quantize the prior estimation of the grayscale image according to the number of classes. Moreover, the matched filter design the coded aperture by computing spectral response between the designed filters and the average training samples. The performance of the proposed classification approach is assessed on two spectral datasets under the OA metric. The simulation test also shows a remarkable improvement in the accuracy classification in up to 99.68% in the OA metric. According to the simulation results, a significant reduction in the number of snapshots in

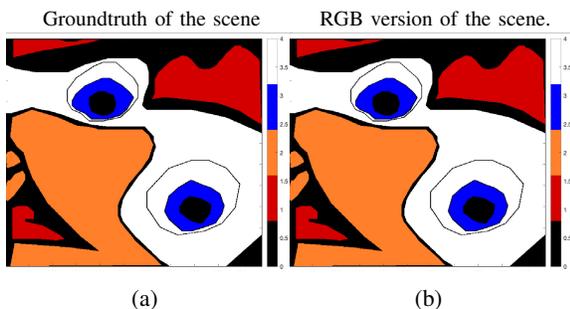


Fig. 10: (a) Groundtruth of the scene, which is composed of $n_c = 4$, and (b) an RGB version of the scene.

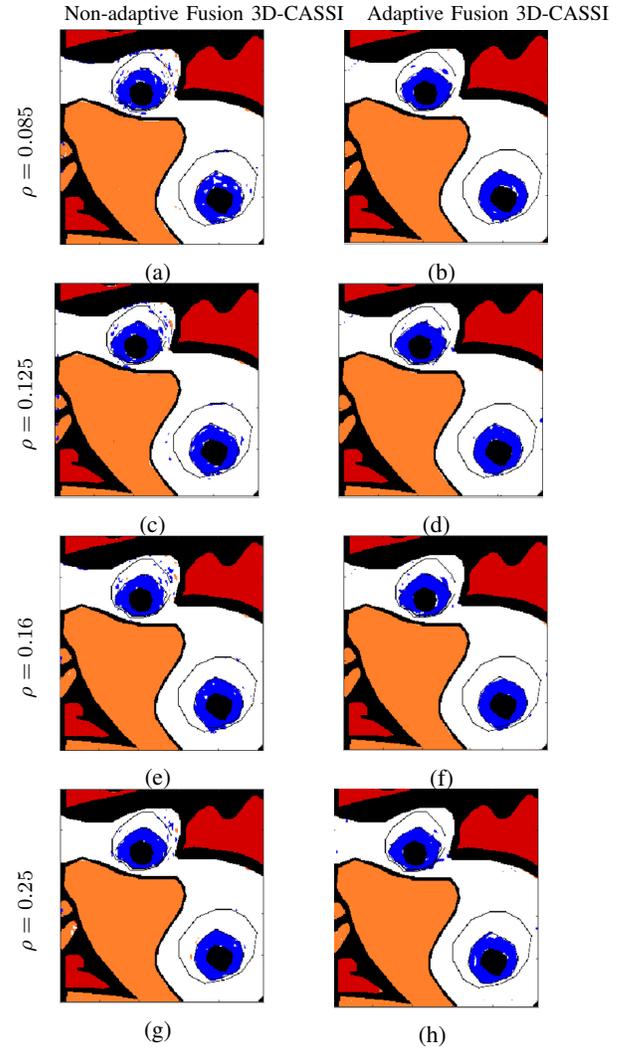


Fig. 11: Classification maps of Hen dataset using non-adaptive coded aperture 3D-CASSI against adaptive coded aperture 3D-CASSI obtained by (a) Non-adaptive with $\rho = 0.085$, AO:97.76%, (b) Adaptive with $\rho = 0.085$, AO:99.19%, (c) Non-adaptive with $\rho = 0.125$, AO:98.10%, (d) Adaptive with $\rho = 0.125$, AO:99.24%, (e) Non-adaptive with $\rho = 0.16$, AO:98.67%, (f) Adaptive with $\rho = 0.16$, AO:99.11%, (g) Non-adaptive with $\rho = 0.25$, AO:99.11%, (h) Adaptive with $\rho = 0.25$, AO:99.19%.

the hyperspectral and multispectral arm is achieved using the proposed approach. Notice that the approach is successfully implemented on real compressive measurements. The testbed implementation shows a significant improvement in the overall accuracy.

FUNDING INFORMATION

Universidad Industrial de Santander under VIE-grant 2699. Departamento Administrativo de Ciencia, Tecnología e Investigación (COLCIENCIAS) under the grant 727 doctorados nacionales 2015. Part of this work was supported by the Air Force Office of Scientific Research (AFOSR) (FA9550-19-1-0293); Agencia Nacional de Investigacion y Desarrollo (ANID FONDECYT) (1181943).

ACKNOWLEDGMENT

Nelson Diaz is supported by Colciencias scholarship. We thank the members of the optics lab from HDSP-UIS for their

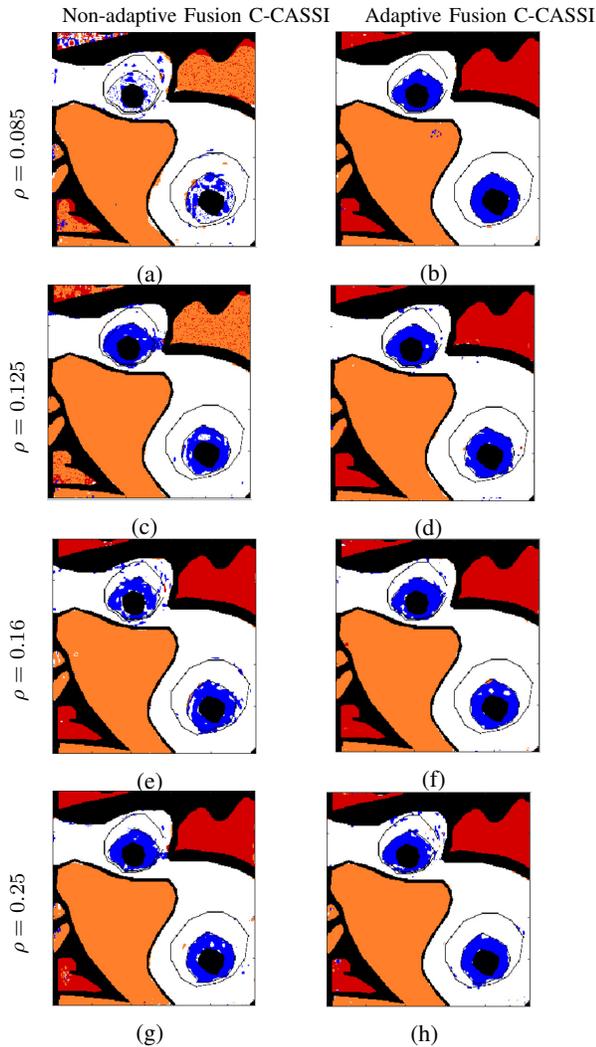


Fig. 12: Classification maps of Hen dataset using non-adaptive coded aperture C-CASSI against adaptive coded aperture C-CASSI obtained by (a) Non-adaptive with $\rho = 0.085$, AO:80.74%, (b) Adaptive with $\rho = 0.085$, AO:99.18%, (c) Non-adaptive with $\rho = 0.125$, AO:84.35%, (d) Adaptive with $\rho = 0.125$, AO:98.6%, (e) Non-adaptive with $\rho = 0.16$, AO:97.03%, (f) Adaptive with $\rho = 0.16$, AO:98.6%, (g) Non-adaptive with $\rho = 0.25$, AO:97.9%, (h) Adaptive with $\rho = 0.25$, AO:98.19%.

help with testbed experiments.

REFERENCES

[1] D. G. Goodenough, A. Dyk, K. O. Niemann, J. S. Pearlman, Hao Chen, T. Han, M. Murdoch, and C. West, "Processing hyperion and ali for forest classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1321–1331, June 2003.

[2] U. Amato, A. Antoniadis, M. F. Carfora, P. Colandrea, V. Cuomo, M. Franzese, S. Pignatti, and C. Serio, "Statistical classification for assessing prisma hyperspectral potential for agricultural land use," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 615–625, April 2013.

[3] C. Yang, J. H. Everitt, Q. Du, B. Luo, and J. Chanussot, "Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 582–592, March 2013.

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

[5] F. Ratle, G. Camps-Valls, and J. Weston, "Semisupervised neural networks for efficient hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2271–2282, May 2010.

[6] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3973–3985, Oct 2011.

[7] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 809–823, March 2012.

[8] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, January 1968.

[9] Z. Ye, H. Li, Y. Song, J. A. Benediktsson, and Y. Y. Tang, "Hyperspectral image classification using principal components-based smooth ordering and multiple 1-d interpolation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 1199–1209, Feb 2017.

[10] N. Falco, J. A. Benediktsson, and L. Bruzzone, "A study on the effectiveness of different independent component analysis algorithms for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2183–2199, June 2014.

[11] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.

[12] J. Arenas-Garcia, K. B. Petersen, G. Camps-Valls, and L. K. Hansen, "Kernel multivariate analysis framework for supervised subspace learning: A tutorial on linear and kernel multivariate methods," *IEEE Signal Processing Magazine*, vol. 30, no. 4, pp. 16–29, 2013.

[13] N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and spatial classification of hyperspectral images based on ica and reduced morphological attribute profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6223–6240, Nov 2015.

[14] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, March 2013.

[15] J. Zhao, Y. Zhong, H. Shu, and L. Zhang, "High-resolution image classification integrating spectral-spatial-location cues by conditional random fields," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4033–4045, Sep. 2016.

[16] X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 95–108, Sep. 2016.

[17] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, Apr 2008. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-47-10-B44>

[18] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express*, vol. 15, no. 21, pp. 14 013–14 027, Oct 2007. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-21-14013>

[19] H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1896–1908, April 2014.

[20] E. Vargas, H. Arguello, and J. Tourneret, "Spectral image fusion from compressive measurements using spectral unmixing and a sparse representation of abundance maps," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 5043–5053, July 2019.

[21] A. Ramirez, H. Arguello, G. R. Arce, and B. M. Sadler, "Spectral image classification from optimal coded-aperture compressive measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3299–3309, June 2014.

[22] J. M. Ramirez and H. Arguello, "Multiresolution compressive feature fusion for spectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9900–9911, Dec 2019.

[23] —, "Spectral image classification from multi-sensor compressive measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 1, pp. 626–636, Jan 2020.

[24] N. Diaz, H. R. Chacon, and H. A. Fuentes, "High-dynamic range compressive spectral imaging by grayscale coded aperture adaptive filtering," *Ingeniería e Investigación*,

vol. 35, no. 3, pp. 53–60, 2015. [Online]. Available: <https://revistas.unal.edu.co/index.php/ingeviv/article/view/49868>

- [25] N. Diaz, C. Hinojosa, and H. Arguello, “Adaptive grayscale compressive spectral imaging using optimal blue noise coding patterns,” *Optics & Laser Technology*, vol. 117, pp. 147 – 157, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0030399218312088>
- [26] N. Diaz, H. Rueda, and H. Arguello, “Adaptive filter design via a gradient thresholding algorithm for compressive spectral imaging,” *Appl. Opt.*, vol. 57, no. 17, pp. 4890–4900, Jun 2018. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-57-17-4890>
- [27] M. Dunlop-Gray, P. Poon, D. Golish, E. Vera, and M. Gehm, “Experimental demonstration of an adaptive architecture for direct spectral imaging classification,” *Opt. Express*, vol. 24, no. 16, pp. 18 307–18 321, Aug 2016. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-24-16-18307>
- [28] Grupo de Inteligencia Computacional, “Hyper Remote Sensing Scenes.” [Online]. Available: http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes
- [29] H. Rueda-Chacon, F. Rojas, D. R. Molina, and H. Arguello, “Demonstration of a compressive hyperspectral image fusion optical imager,” in *Imaging and Applied Optics Congress*. Optical Society of America, 2020, p. CW4B.6. [Online]. Available: <http://www.osapublishing.org/abstract.cfm?URI=COSI-2020-CW4B.6>



Esteban Vera (M’98) received the B.S. degree and Engineering diploma in electronics engineering, and the M.S. and Ph.D. degrees in electrical engineering from Universidad de Concepcion, Concepcion, Chile, in 1996, 1999, 2003, and 2010, respectively. From 2001 to 2007, he worked as an electronics engineer for different large telescope projects, first with the Paranal Observatory and then with the Gemini Observatory. In 2010, he was a post-doctoral research associate with the University of Arizona, Tucson, AZ, USA, and in 2013, he became a research scientist at Duke University, Durham, NC, USA. In 2016, he joined the School of Electrical Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, where he is currently an associate professor. His research interests span the fields of computational imaging, compressed sensing, inverse problems, signal processing, machine learning, and the development of scientific and astronomical instrumentation. He is a member of the SPIE and a senior member of the OSA.



Nelson Diaz (S’19–M’21) received the B.Sc.(Eng.) in computer science engineering degree and the M.Sc. degree in computer science from the Universidad Industrial de Santander, Colombia, in 2012 and 2015, respectively. He received a Ph.D. in Engineering in 2020 from the Universidad Industrial de Santander, Colombia. Currently, he holds a post-doctoral position in Pontificia Universidad Católica de Valparaíso (PUCV), Chile under the supervision of Profesor Esteban Vera. His research interests include sparse image representation, adaptive sensing,

and spectral image classification.



Henry Arguello (S’11–M’13–SM’17) received the B.Sc. Eng. degree in electrical engineering and the M.Sc. degree in electrical power from the Universidad Industrial de Santander, Bucaramanga, Colombia, in 2000 and 2003, respectively, and the Ph.D. degree in electrical engineering from the University of Delaware, Newark, DE, USA, in 2013. He is currently an Associate Professor with the Department of Systems Engineering, Universidad Industrial de Santander. In first semester 2020, he was a Visiting Professor with Stanford University, Stanford, CA,

USA, funded by Fulbright. His research interests include high-dimensional signal processing, optical imaging, compressed sensing, hyperspectral imaging, and CI.



Juan Ramirez (M’19) received the B.S. degree in electrical engineering, the M.S. degree in biomedical engineering, and the Ph.D. degree in applied sciences from the Universidad de Los Andes, Mérida, Venezuela, in 2002, 2007 and 2017, respectively. He has been a Professor with the Electrical Engineering Department, Universidad de Los Andes, Venezuela, from 2004 to 2019. He is currently a Post-Doctoral Intern with the Department of Computer Science, Universidad Industrial de Santander, Bucaramanga, Colombia. His research interests include robust

sparse signal processing, robust sparse signal representation, and spectral image classification.